

File:../data/experiments/base_cased/training_folder/
no_attention/predictions_strong_matching/aida_train
checkpoint=54, opt_thr=0.03, model=local:
pem_similarityglobal: local

Sample: b'1_EU'

EU rejects [1 1 German 1] call to boycott [2 2 British 2] lamb . Peter Blackburn [3 3 BRUSSELS 3] 1996-08-22 The [4 4 European Commission 4] said on Thursday it disagreed with [5 5 German 5] advice to consumers to shun [6 6 British 6] lamb until scientists determine whether mad cow disease can be transmitted to sheep . [7 7 Germany 7] 's representative to the [8 8 European Union 8] 's veterinary committee Werner Zwingmann said on Wednesday consumers should buy sheepmeat from countries other than [9 9 Britain 9] until the scientific advice was clearer . " We do n't support any such recommendation because we do n't see any grounds for it , " the [10 10 Commission 10] 's chief spokesman Nikolaus van der Pas told a news briefing . He said further scientific study was required and if it was found that action was needed it should be taken by the [11 11 European Union 11] . He said a proposal last month by EU Farm Commissioner [12 12 Franz Fischler 12] to ban sheep brains , spleens and spinal cords from the human and animal food chains was a highly specific and precautionary move to protect human health . Fischler proposed EU-wide measures after reports from [13 13 Britain 13] and [14 14 France 14] that under laboratory conditions sheep could contract Bovine Spongiform Encephalopathy ([15 15 BSE 15]) -- mad cow disease . But Fischler agreed to review his proposal after the EU 's standing veterinary committee , national animal health officials , questioned if such action was justified as there was only a slight risk to human health . [16 16 Spanish 16] Farm Minister [17 17 Loyola de Palacio 17] had earlier accused Fischler at an EU farm ministers ' meeting of causing unjustified alarm through " dangerous generalisation . " . Only [18 18 France 18] and [19 19 Britain 19] backed Fischler 's proposal . The EU 's scientific veterinary and multidisciplinary committees are due to re-examine the issue early next month and make recommendations to the senior veterinary officials . Sheep have long been known to contract scrapie , a brain-wasting disease similar to [20 20 BSE 20] which is believed to have been transferred to cattle through feed containing animal waste . [21 21 British 21] farmers denied on Thursday there was any danger to human health from their sheep , but expressed concern that [22 22 German 22] government advice to consumers to avoid [23 23 British 23] lamb might influence consumers across [24 24 Europe 24] . " What we have to be extremely careful of is how other countries are going to take [25 25 Germany 25] 's lead , " Welsh National Farmers ' Union (NFU) chairman John Lloyd Jones said on BBC radio . [26 26 B

on [26] has led efforts to protect public health after consumer confidence collapsed in March after a [27 27 British [27] report suggested humans could contract an illness similar to mad cow disease by eating contaminated beef . [28 28 Germany [28] imported 47,600 sheep from [29 29 Britain [29] last year , nearly half of total imports . It brought in 4,275 tonnes of [30 30 British [30] mutton , some 10 percent of overall imports .

1: 11867 Germany
2: 31717 United_Kingdom
3: 3708 Brussels
4: 9974 European_Commission
5: 11867 Germany
6: 31717 United_Kingdom
7: 11867 Germany
8: 9317 European_Union
9: 31717 United_Kingdom
10: 9974 European_Commission
11: 9317 European_Union
12: 626779 Franz_Fischler
13: 31717 United_Kingdom
14: 5843419 France
15: 19344418 Bovine_spongiform_encephalopathy
16: 26667 Spain
17: 6394317 Loyola_de_Palacio
18: 5843419 France
19: 31717 United_Kingdom
20: 19344418 Bovine_spongiform_encephalopathy
21: 31717 United_Kingdom
22: 11867 Germany
23: 31717 United_Kingdom
24: 9239 Europe
25: 11867 Germany
26: 3295 Bonn
27: 31717 United_Kingdom
28: 11867 Germany
29: 31717 United_Kingdom
30: 31717 United_Kingdom
1: 11867 Germany
[pem_pos=1, final_scores=0.7692233,
final_local_scores=0.7562699, pem=-1.0465823,
similarity=32.349792]
2: 31717 United_Kingdom
[pem_pos=0, final_scores=0.60063016,
final_local_scores=0.5542209, pem=-0.48936236,
similarity=23.289999]
3: 3708 Brussels
[pem_pos=0, final_scores=0.5661251,
final_local_scores=0.5611217, pem=-0.025440974,
similarity=20.973057]
4: 9974 European_Commission
[pem_pos=0, final_scores=0.4685635,

```
final_local_scores=0.48390782, pem=0.0, similarity=18.530487]
5: 11867 Germany
    [pem_pos=1, final_scores=0.79953974,
final_local_scores=0.789386, pem=-1.0465823,
similarity=33.338043]
6: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.57615966,
final_local_scores=0.5274906, pem=-0.48936236,
similarity=22.492313]
7: 11867 Germany
    [pem_pos=0, final_scores=0.7093013,
final_local_scores=0.69081426, pem=-0.15529114,
similarity=25.549486]
8: 9317 European_Union
    [pem_pos=0, final_scores=0.4802664,
final_local_scores=0.46786928, pem=-0.005814796,
similarity=18.083488]
9: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.59469193,
final_local_scores=0.5477343, pem=-0.61114657,
similarity=23.758707]
10: 9974 European_Commission
    [pem_pos=0, final_scores=0.5964151,
final_local_scores=0.6235662, pem=-0.76247066,
similarity=26.844606]
11: 9317 European_Union
    [pem_pos=0, final_scores=0.47269124,
final_local_scores=0.45959455, pem=-0.005814796,
similarity=17.836554]
12: 626779 Franz_Fischler
    [pem_pos=0, final_scores=0.5718493,
final_local_scores=0.73493683, pem=0.0, similarity=26.021694]
13: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.6034721,
final_local_scores=0.5573253, pem=-0.61114657,
similarity=24.04492]
14: 5843419 France
    [pem_pos=0, final_scores=0.7177666,
final_local_scores=0.707847, pem=-0.16240834,
similarity=26.09648]
15: 19344418 Bovine_spongiform_encephalopathy
    [pem_pos=0, final_scores=0.58326113,
final_local_scores=0.6922914, pem=-0.2201173,
similarity=25.9461]
16: 26667 Spain
    [pem_pos=1, final_scores=0.63891596,
final_local_scores=0.6491298, pem=-1.2425895,
similarity=30.218437]
17: 6394317 Loyola_de_Palacio
    [pem_pos=0, final_scores=0.577029,
final_local_scores=0.71087813, pem=0.0, similarity=25.303734]
18: 5843419 France
```

```
[pem_pos=0, final_scores=0.6465724,
final_local_scores=0.6300782, pem=-0.16240834,
similarity=23.775703]
19: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5746789,
final_local_scores=0.5258731, pem=-0.61114657,
similarity=23.106327]
20: 19344418 Bovine_spongiform_encephalopathy
    [pem_pos=0, final_scores=0.6973525,
final_local_scores=0.8169189, pem=-0.2201173,
similarity=29.665234]
21: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5080914,
final_local_scores=0.45313638, pem=-0.48936236,
similarity=20.273436]
22: 11867 Germany
    [pem_pos=1, final_scores=0.7989864,
final_local_scores=0.7887816, pem=-1.0465823,
similarity=33.320007]
23: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5567861,
final_local_scores=0.506328, pem=-0.48936236,
similarity=21.860779]
24: 9239 Europe
    [pem_pos=0, final_scores=0.6433383,
final_local_scores=0.63881993, pem=-0.086338006,
similarity=23.622894]
25: 11867 Germany
    [pem_pos=0, final_scores=0.7885043,
final_local_scores=0.77733153, pem=-0.15529114,
similarity=28.131334]
26: 3295 Bonn
    [pem_pos=0, final_scores=0.35958666,
final_local_scores=0.34818336, pem=0.0, similarity=14.480199]
27: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.55382544,
final_local_scores=0.5030939, pem=-0.48936236,
similarity=21.764267]
28: 11867 Germany
    [pem_pos=0, final_scores=0.7550072,
final_local_scores=0.74074095, pem=-0.15529114,
similarity=27.039398]
29: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.58065206,
final_local_scores=0.5323979, pem=-0.61114657,
similarity=23.30104]
30: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.55434734,
final_local_scores=0.50366396, pem=-0.48936236,
similarity=21.781279]
```

Sample: b'2_Rare'

Rare [1 1 Jimi Hendrix 1] song draft sells for almost \$ 17,000 . [2 2 LONDON 2] 1996-08-22 A rare early handwritten draft of a song by [3 3 U.S. 3] guitar legend [4 4 Jimi Hendrix 4] was sold for almost \$ 17,000 on Thursday at an auction of some of the late musician 's favourite possessions . A Florida restaurant paid 10,925 pounds (\$ 16,935) for the draft of " Ai n't no telling " , which [5 5 Hendrix 5] penned on a piece of [6 6 London 6] hotel stationery in late 1966 . At the end of a January 1967 concert in the [7 7 English 7] city of [8 8 Nottingham 8] he threw the sheet of paper into the audience , where it was retrieved by a fan . Buyers also snapped up 16 other items that were put up for auction by [9 9 Hendrix 9] 's former girlfriend Kathy Etchingham , who lived with him from 1966 to 1969 . They included a black lacquer and mother of pearl inlaid box used by [10 10 Hendrix 10] to store his drugs , which an anonymous [11 11 Australian 11] purchaser bought for 5,060 pounds (\$ 7,845) . The guitarist died of a drugs overdose in 1970 aged 27 .

```
1: 16095 Jimi_Hendrix
2: 17867 London
3: 3434750 United_States
4: 16095 Jimi_Hendrix
5: 16095 Jimi_Hendrix
6: 17867 London
7: 9316 England
8: 39470 Nottingham
9: 16095 Jimi_Hendrix
10: 16095 Jimi_Hendrix
11: 4689264 Australia
1: 16095 Jimi_Hendrix
    [pem_pos=0, final_scores=0.577522,
final_local_scores=0.53772557, pem=0.0, similarity=20.136517]
2: 17867 London
    [pem_pos=0, final_scores=0.621033,
final_local_scores=0.6431907, pem=-0.09086737,
similarity=23.777956]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.47617614,
final_local_scores=0.512659, pem=-0.14075686,
similarity=20.153938]
4: 16095 Jimi_Hendrix
    [pem_pos=0, final_scores=0.45069236,
final_local_scores=0.3938164, pem=0.0, similarity=15.841981]
5: 16095 Jimi_Hendrix
    [pem_pos=0, final_scores=0.58709735,
final_local_scores=0.54818517, pem=0.0, similarity=20.448652]
6: 17867 London
    [pem_pos=0, final_scores=0.48028362,
final_local_scores=0.48944336, pem=-0.09086737,
similarity=19.189829]
7: 9316 England
    [pem_pos=1, final_scores=0.49976128,
```

```
final_local_scores=0.53363055, pem=-1.3527626,  
similarity=27.370846]  
8: 39470 Nottingham  
    [pem_pos=0, final_scores=0.6378567,  
final_local_scores=0.6871477, pem=-0.29126778,  
similarity=26.17953]  
9: 16095 Jimi_Hendrix  
    [pem_pos=0, final_scores=0.614175,  
final_local_scores=0.5777635, pem=0.0, similarity=21.331327]  
10: 16095 Jimi_Hendrix  
    [pem_pos=0, final_scores=0.5920033,  
final_local_scores=0.5535443, pem=0.0, similarity=20.608578]  
11: 4689264 Australia  
    [pem_pos=0, final_scores=0.45074314,  
final_local_scores=0.5193654, pem=-0.491009,  
similarity=22.258795]
```

Sample: b'3_China'

[1 1 China 1] says [2 2 Taiwan 2] spoils atmosphere for talks . [3 3 BEIJING 3] 1996-08-22 [4 4 China 4] on Thursday accused [5 5 Taipei 5] of spoiling the atmosphere for a resumption of talks across the [6 6 Taiwan Strait 6] with a visit to [7 7 Ukraine 7] by [8 8 Taiwanese 8] Vice President [9 9 Lien Chan 9] this week that infuriated [10 10 Beijing 10] . Speaking only hours after [11 11 Chinese 11] state media said the time was right to engage in political talks with [12 12 Taiwan 12] , [13 13 Foreign Ministry 13] spokesman Shen Guofang told [1 Reuters 1] : " The necessary atmosphere for the opening of the talks has been disrupted by the [14 14 Taiwan 14] authorities . " State media quoted [15 15 China 15] 's top negotiator with [16 16 Taipei 16] , Tang Shubei , as telling a visiting group from [17 17 Taiwan 17] on Wednesday that it was time for the rivals to hold political talks . " Now is the time for the two sides to engage in political talks ... that is to end the state of hostility , " Thursday 's overseas edition of the People 's Daily quoted Tang as saying . The foreign ministry 's Shen told [18 18 Reuters Television 18] in an interview he had read reports of Tang 's comments but gave no details of why the negotiator had considered the time right for talks with [19 19 Taiwan 19] , which [20 20 Beijing 20] considers a renegade province . [21 21 China 21] , which has long opposed all [22 22 Taipei 22] efforts to gain greater international recognition , was infuriated by a visit to [23 23 Ukraine 23] this week by [24 24 Taiwanese 24] Vice President Lien .

```
1: 5405 China  
2: 25734 Taiwan  
3: 18603746 Beijing  
4: 5405 China  
5: 57648 Taipei  
6: 91052 Taiwan_Strait  
7: 31750 Ukraine
```

8: 25734 Taiwan
9: 140551 Lien_Chan
10: 18603746 Beijing
11: 5405 China
12: 25734 Taiwan
13: 7453497
Ministry_of_Foreign_Affairs_of_the_People's_Republic_of_China
14: 25734 Taiwan
15: 5405 China
16: 57648 Taipei
17: 25734 Taiwan
18: 18998750 Reuters
19: 25734 Taiwan
20: 18603746 Beijing
21: 5405 China
22: 57648 Taipei
23: 31750 Ukraine
24: 25734 Taiwan
1: 5405 China
[pem_pos=0, final_scores=0.56309956,
final_local_scores=0.51305944, pem=-0.15669827,
similarity=20.25258]
2: 25734 Taiwan
[pem_pos=0, final_scores=0.54182214,
final_local_scores=0.48664397, pem=-0.077886656,
similarity=19.0357]
3: 18603746 Beijing
[pem_pos=0, final_scores=0.49783403,
final_local_scores=0.47866172, pem=-0.042161036,
similarity=18.603212]
4: 5405 China
[pem_pos=0, final_scores=0.6087825,
final_local_scores=0.56296104, pem=-0.15669827,
similarity=21.741743]
5: 57648 Taipei
[pem_pos=0, final_scores=0.37298518,
final_local_scores=0.33578026, pem=0.0, similarity=14.110065]
6: 91052 Taiwan_Strait
[pem_pos=0, final_scores=0.5116962,
final_local_scores=0.5674238, pem=0.0, similarity=21.022772]
7: 31750 Ukraine
[pem_pos=0, final_scores=0.6320607,
final_local_scores=0.6511578, pem=-0.35407752,
similarity=25.447088]
8: 25734 Taiwan
[pem_pos=0, final_scores=0.6930645,
final_local_scores=0.65185326, pem=-0.76474255,
similarity=27.701103]
9: 140551 Lien_Chan
[pem_pos=0, final_scores=0.38345897,
final_local_scores=0.4399405, pem=0.0, similarity=17.218416]
10: 18603746 Beijing

```
[pem_pos=0, final_scores=0.27319473,
final_local_scores=0.23327741, pem=-0.042161036,
similarity=11.280455]
11: 5405 China
    [pem_pos=1, final_scores=0.71000844,
final_local_scores=0.67353505, pem=-1.4196125,
similarity=31.909416]
12: 25734 Taiwan
    [pem_pos=0, final_scores=0.44818777,
final_local_scores=0.38436264, pem=-0.077886656,
similarity=15.983421]
13: 7453497
Ministry_of_Foreign_Affairs_of_the_People's_Republic_of_China
    [pem_pos=2, final_scores=0.17036994,
final_local_scores=0.18782213, pem=-2.5268707,
similarity=23.436216]
14: 25734 Taiwan
    [pem_pos=0, final_scores=0.50581795,
final_local_scores=0.44731486, pem=-0.077886656,
similarity=17.86204]
15: 5405 China
    [pem_pos=0, final_scores=0.6168894,
final_local_scores=0.5718166, pem=-0.15669827,
similarity=22.006012]
16: 57648 Taipei
    [pem_pos=0, final_scores=0.4273947,
final_local_scores=0.39521438, pem=0.0, similarity=15.883698]
17: 25734 Taiwan
    [pem_pos=0, final_scores=0.48893934,
final_local_scores=0.42887753, pem=-0.077886656,
similarity=17.311834]
18: 18998750 Reuters
    [pem_pos=0, final_scores=0.5671279,
final_local_scores=0.6472673, pem=0.0, similarity=23.40546]
19: 25734 Taiwan
    [pem_pos=0, final_scores=0.46424782,
final_local_scores=0.40190583, pem=-0.077886656,
similarity=16.506945]
20: 18603746 Beijing
    [pem_pos=0, final_scores=0.38042074,
final_local_scores=0.35040554, pem=-0.042161036,
similarity=14.775791]
21: 5405 China
    [pem_pos=0, final_scores=0.56073815,
final_local_scores=0.5104799, pem=-0.15669827,
similarity=20.175602]
22: 57648 Taipei
    [pem_pos=0, final_scores=0.36352044,
final_local_scores=0.32544148, pem=0.0, similarity=13.801536]
23: 31750 Ukraine
    [pem_pos=0, final_scores=0.54351246,
final_local_scores=0.5544323, pem=-0.35407752,
```

```
similarity=22.560606]
24: 25734 Taiwan
    [pem_pos=0, final_scores=0.6406142,
final_local_scores=0.5945594, pem=-0.76474255,
similarity=25.991339]
1: 18998750 Reuters
    [pem_pos=0, final_scores=0.17860226,
final_local_scores=0.222862, pem=0.0, similarity=10.740359]

Sample: b'4_China'
[1 1 China 1] says time right for [2 2 Taiwan 2] talks . [3 3
BEIJING 3] 1996-08-22 [4 4 China 4] has said it was time for
political talks with [5 5 Taiwan 5] and that the rival island
should take practical steps towards that goal . Consultations
should be held to set the time and format of the talks , the
official [6 6 Xinhua 6] news agency quoted Tang Shubei ,
executive vice chairman of the [7 7 Association for Relations
Across the Taiwan Straits 7] , as saying late on Wednesday .
1: 5405 China
2: 25734 Taiwan
3: 18603746 Beijing
4: 5405 China
5: 25734 Taiwan
6: 263168 Xinhua_News_Agency
7: 921919 Association_for_Relations_Across_the_Taiwan_Straits
1: 5405 China
    [pem_pos=0, final_scores=0.5812792,
final_local_scores=0.5384426, pem=-0.15669827,
similarity=21.010065]
2: 25734 Taiwan
    [pem_pos=0, final_scores=0.5094599,
final_local_scores=0.4619779, pem=-0.077886656,
similarity=18.299614]
3: 18603746 Beijing
    [pem_pos=0, final_scores=0.47395986,
final_local_scores=0.47649086, pem=-0.042161036,
similarity=18.53843]
4: 5405 China
    [pem_pos=0, final_scores=0.5831788,
final_local_scores=0.5405176, pem=-0.15669827,
similarity=21.071987]
5: 25734 Taiwan
    [pem_pos=0, final_scores=0.43917936,
final_local_scores=0.385207, pem=-0.077886656,
similarity=16.008617]
6: 263168 Xinhua_News_Agency
    [pem_pos=0, final_scores=0.47784573,
final_local_scores=0.5093357, pem=0.0, similarity=19.289307]
7: 921919 Association_for_Relations_Across_the_Taiwan_Straits
    [pem_pos=0, final_scores=0.5285737,
final_local_scores=0.6296154, pem=0.0, similarity=22.878693]
```

Sample: b'5_German'

[1 1 German 1] July car registrations up 14.2 pct yr / yr . [2 2 FRANKFURT 2] 1996-08-22 [3 3 German 3] first-time registrations of motor vehicles jumped 14.2 percent in July this year from the year-earlier period , the Federal office for motor vehicles said on Thursday . The office said 356,725 new cars were registered in July 1996 -- 304,850 passenger cars and 15,613 trucks . The figures represent a 13.6 percent increase for passenger cars and a 2.2 percent decline for trucks from July 1995 . Motor-bike registration rose 32.7 percent in the period . The growth was partly due to an increased number of [4 4 Germans 4] buying [5 5 German 5] cars abroad , while manufacturers said that domestic demand was weak , the federal office said . Almost all [6 6 German 6] car manufacturers posted gains in registration numbers in the period . [7 7 Volkswagen AG 7] won 77,719 registrations , slightly more than a quarter of the total . [8 8 Opel AG 8] together with [9 9 General Motors 9] came in second place with 49,269 registrations , 16.4 percent of the overall figure . Third was [10 10 Ford 10] with 35,563 registrations , or 11.7 percent . Only Seat and [11 11 Porsche 11] had fewer registrations in July 1996 compared to last year 's July . Seat posted 3,420 registrations compared with 5522 registrations in July a year earlier . [12 12 Porsche 12] 's registrations fell to 554 from 643 .

1: 11867 Germany

2: 10992 Frankfurt

3: 11867 Germany

4: 11867 Germany

5: 11867 Germany

6: 11867 Germany

7: 32413 Volkswagen

8: 22284 Opel

9: 12102 General_Motors

10: 30433662 Ford_Motor_Company

11: 24365 Porsche

12: 24365 Porsche

1: 11867 Germany

[pem_pos=1, final_scores=0.7287017,
final_local_scores=0.72654516, pem=-1.0465823,
similarity=31.46275]

2: 10992 Frankfurt

[pem_pos=0, final_scores=0.5119963,
final_local_scores=0.51202464, pem=-0.07062824,
similarity=19.753637]

3: 11867 Germany

[pem_pos=1, final_scores=0.64911926,
final_local_scores=0.63961345, pem=-1.0465823,
similarity=28.868532]

4: 11867 Germany

[pem_pos=1, final_scores=0.61259896,
final_local_scores=0.59972054, pem=-1.5806074,

```
similarity=30.58216]
5: 11867 Germany
    [pem_pos=1, final_scores=0.8252426,
final_local_scores=0.8320014, pem=-1.0465823,
similarity=34.60977]
6: 11867 Germany
    [pem_pos=1, final_scores=0.77154547,
final_local_scores=0.7733454, pem=-1.0465823,
similarity=32.85936]
7: 32413 Volkswagen
    [pem_pos=1, final_scores=0.1255612,
final_local_scores=0.065562755, pem=-0.61542445,
similarity=9.39301]
8: 22284 Opel
    [pem_pos=0, final_scores=0.3444882,
final_local_scores=0.32538176, pem=0.0, similarity=13.799753]
9: 12102 General_Motors
    [pem_pos=0, final_scores=0.3339314,
final_local_scores=0.3179448, pem=0.0, similarity=13.57782]
10: 30433662 Ford_Motor_Company
    [pem_pos=0, final_scores=0.46527022,
final_local_scores=0.4552393, pem=-0.3393453,
similarity=19.520374]
11: 24365 Porsche
    [pem_pos=0, final_scores=0.34465444,
final_local_scores=0.32587546, pem=-0.0039392714,
similarity=13.835909]
12: 24365 Porsche
    [pem_pos=0, final_scores=0.39321357,
final_local_scores=0.37891895, pem=-0.0039392714,
similarity=15.418833]
```

Sample: b'6_GREEK'

```
[1 1 GREEK 1] SOCIALISTS GIVE GREEN LIGHT TO PM FOR
ELECTIONS . [2 2 ATHENS 2] 1996-08-22 The [3 3 Greek 3]
socialist party 's executive bureau gave the green light to
Prime Minister [4 4 Costas Simitis 4] to call snap
elections , its general secretary [5 5 Costas Skandalidis 5]
told reporters . Prime Minister [6 6 Costas Simitis 6] is
going to make an official announcement after a cabinet
meeting later on Thursday , said Skandalidis . -- Dimitris
Kontogiannis , Athens Newsroom +301 3311812-4
```

```
1: 12108 Greece
2: 1216 Athens
3: 12108 Greece
4: 393824 Costas_Simitis
5: 13871790 Kostas_Skandalidis
6: 393824 Costas_Simitis
1: 12108 Greece
    [pem_pos=1, final_scores=0.40298563,
final_local_scores=0.38155353, pem=-2.3247876,
similarity=28.118587]
```

```
2: 1216 Athens
    [pem_pos=0, final_scores=0.47133684,
final_local_scores=0.4843186, pem=-0.28831622,
similarity=20.110655]
3: 12108 Greece
    [pem_pos=1, final_scores=0.75398666,
final_local_scores=0.7649689, pem=-2.3247876,
similarity=39.560467]
4: 393824 Costas_Simitis
    [pem_pos=0, final_scores=0.6821422,
final_local_scores=0.71476763, pem=0.0, similarity=25.419806]
5: 13871790 Kostas_Skandalidis
    [pem_pos=0, final_scores=0.48898202,
final_local_scores=0.57082933, pem=0.0, similarity=21.1244]
6: 393824 Costas_Simitis
    [pem_pos=0, final_scores=0.6411094,
final_local_scores=0.6699456, pem=0.0, similarity=24.082226]
```

Sample: b'7_BayerVB'

BayerVB sets C\$ 100 million six-year bond . [1 1 LONDON 1]
1996-08-22 The following bond was announced by lead manager
Toronto Dominion . BORROWER [2 2 BAYERISCHE VEREINSBANK 2]
AMT C\$ 100 MLN COUPON 6.625 MATURITY 24.SEP.02 TYPE STRAIGHT
ISS PRICE 100.92 PAY DATE 24.SEP.96 FULL FEES 1.875 REOFFER
99.32 SPREAD +20 BP MOODY AA1 LISTING LUX PAY FREQ = S&P =
DENOMS (K) 1-10-100 SALE LIMITS US / UK / CA NEG PLG NO CRS
DEFLT NO FORCE MAJ = GOV LAW [3 3 GERMAN 3] HOME CTRY = TAX
PROVS STANDARD MGT / UND 0.275 SELL CONC 1.60 PRAECIP =
UNDERLYING GOVT BOND 7.0 PCT SEPT 2001 NOTES [4 4 BAYERISCHE V
EREINSBANK 4] IS JOINT LEAD MANAGER -- London Newsroom +44
171 542 7658

```
1: 17867 London
2: 2092024 HypoVereinsbank
3: 11867 Germany
4: 2092024 HypoVereinsbank
1: 17867 London
    [pem_pos=0, final_scores=0.5425591,
final_local_scores=0.6461557, pem=-0.09086737,
similarity=23.866438]
2: 2092024 HypoVereinsbank
    [pem_pos=0, final_scores=0.55434674,
final_local_scores=0.58737004, pem=0.0, similarity=21.618006]
3: 11867 Germany
    [pem_pos=1, final_scores=0.7293918,
final_local_scores=0.81044585, pem=-1.0465823,
similarity=33.96651]
4: 2092024 HypoVereinsbank
    [pem_pos=0, final_scores=0.5047816,
final_local_scores=0.5332276, pem=0.0, similarity=20.002289]
```

Sample: b'8_Venantius'

Venantius sets \$ 300 million January 1999 FRN . [1 1 LONDON 1

] 1996-08-22 The following floating-rate issue was announced by lead manager Lehman Brothers International . BORROWER VENANTIUS AB ([2 2 SWEDISH 2] NATIONAL MORTGAGE AGENCY) AMT \$ 300 MLN SPREAD - 12.5 BP MATURITY 21.JAN.99 TYPE FRN BASE 3M LIBOR PAY DATE S23.SEP.96 LAST MOODY AA3 ISS PRICE 99.956 FULL FEES 10 BP LAST S&P AA+ REOFFER = NOTES S SHORT FIRST COUPON LISTING [3 3 LONDON 3] DENOMS (K) 1-10-100 SALE LIMITS US / UK / JP / FR NEG PLG YES CRS DEFLT NO FORCE MAJ IPMA 2 GOV LAW [4 4 ENGLISH 4] HOME CTRY [5 5 SWEDEN 5] TAX PROVS STANDARD MGT / UND 5 BP SELL CONC 5 BP PRAECIP = NOTES ISSUED OFF EMTN PROGRAMME -- London Newsroom +44 171 542 8863

1: 17867 London
2: 5058739 Sweden
3: 17867 London
4: 9316 England
5: 5058739 Sweden

1: 17867 London
[pem_pos=0, final_scores=0.7328295,
final_local_scores=0.708213, pem=-0.09086737,
similarity=25.718351]
2: 5058739 Sweden
[pem_pos=0, final_scores=0.5548098,
final_local_scores=0.5515179, pem=-0.63135606,
similarity=23.98152]
3: 17867 London
[pem_pos=0, final_scores=0.61497897,
final_local_scores=0.5794791, pem=-0.09086737,
similarity=21.876675]
4: 9316 England
[pem_pos=1, final_scores=0.46480113,
final_local_scores=0.4254319, pem=-1.3527626,
similarity=24.141983]
5: 5058739 Sweden
[pem_pos=0, final_scores=0.5216348,
final_local_scores=0.5152793, pem=-0.2135508,
similarity=20.627996]

Sample: b'9_Port'
Port conditions update - [1 1 Syria 1] - Lloyds Shipping .
Port conditions from Lloyds Shipping Intelligence Service --
LATTAKIA , Aug 10 - waiting time at Lattakia and [2 2 Tartus
2] presently 24 hours .

1: 7515849 Syria
2: 879212 Tartus
1: 7515849 Syria
[pem_pos=0, final_scores=0.4703949,
final_local_scores=0.50070983, pem=-0.26162508,
similarity=20.45465]
2: 879212 Tartus
[pem_pos=0, final_scores=0.3619359,
final_local_scores=0.38223487, pem=0.0, similarity=15.496365]

Sample: b'10_Israel'

[1 1 Israel 1] plays down fears of war with [2 2 Syria 2] . Colleen Siegel [3 3 JERUSALEM 3] 1996-08-22 [4 4 Israel 4] 's outgoing peace negotiator with [5 5 Syria 5] said on Thursday current tensions between the two countries appeared to be a storm in a teacup . [6 6 Itamar Rabinovich 6] , who as [7 7 Israel 7] 's ambassador to [8 8 Washington 8] conducted unfruitful negotiations with [9 9 Syria 9] , told [10 10 Israel Radio 10] it looked like [11 11 Damascus 11] wanted to talk rather than fight . " It appears to me the [12 12 Syrian 12] priority is still to negotiate . The [13 13 Syrians 13] are confused , they are definitely tense , but the general assessment here in [14 14 Washington 14] is that this is essentially a storm in a teacup , " he said . [15 15 Rabinovich 15] is winding up his term as ambassador . He will be replaced by [16 16 Eliahu Ben-Elissar 16] , a former [17 17 Israeli 17] envoy to [18 18 Egypt 18] and right-wing [19 19 Likud 19] party politician . [20 20 Israel 20] on Wednesday sent [21 21 Syria 21] a message , via 1 Washington , saying it was committed to peace and wanted to open negotiations without preconditions . But it slammed [22 22 Damascus 22] for creating what it called a dangerous atmosphere . [23 23 Syria 23] accused [24 24 Israel 24] on Wednesday of launching a hysterical campaign against it after [25 25 Israeli 25] television reported that [26 26 Damascus 26] had recently test fired a missile . It said its arms purchases were for defensive purposes . " The message that we sent to ([27 27 Syrian 27] President Hafez al-) [1 2 Assad 1] is that [28 28 Israel 28] is ready at any time without preconditions to enter peace negotiations , " [29 29 Israeli 29] Foreign Minister [30 30 David Levy 30] told [31 31 Israel Radio 31] in an interview . Tension has mounted since [32 32 Israeli 32] Prime Minister [33 33 Benjamin Netanyahu 33] took office in June vowing to retain the [34 34 Golan Heights 34] [35 35 Israel 35] captured from [36 36 Syria 36] in the 1967 [37 37 Middle East 37] war . Israeli-Syrian peace talks have been deadlocked over the [38 38 Golan 38] since 1991 despite the previous government 's willingness to make [39 39 Golan 39] concessions . Peace talks between the two sides were last held in February . " The voices coming out of [40 40 Damascus 40] are bad , not good . The media ... are full of expressions and declarations that must be worrying ... this artificial atmosphere is very dangerous because those who spread it could become its prisoners , " Levy said . " We expect from [41 41 Syria 41] , if its face is to peace , that it will answer [42 42 Israel 42] 's message to enter peace negotiations because that is our goal , " he said . " We do not want a war , [43 43 God 43] forbid . No one benefits from wars . " [44 44 Israel 44] 's [45 45 Channel Two 45] television said [46 46 Damascus 46] had sent a " calming signal " to [47 47 Israel 47] . It gave no source for the report . [48 48 Netanyahu 48] and Levy 's spokesmen said they

could not confirm it . The television also said that [49 49 Netanyahu 49] had sent messages to reassure [50 50 Syria 50] via Cairo , the [51 51 United States 51] and 3 Moscow .

1: 9282173 Israel
2: 7515849 Syria
3: 16043 Jerusalem
4: 9282173 Israel
5: 7515849 Syria
6: 17944843 Itamar_Rabinovich
7: 9282173 Israel
8: 108956 Washington,_D.C.
9: 7515849 Syria
10: 2820017 Kol_Yisrael
11: 7515849 Syria
12: 7515849 Syria
13: 7515849 Syria
14: 108956 Washington,_D.C.
15: 17944843 Itamar_Rabinovich
16: 8320263 Eliyahu_Ben-Elissar
17: 9282173 Israel
18: 8087628 Egypt
19: 206984 Likud
20: 9282173 Israel
21: 7515849 Syria
22: 7515849 Syria
23: 7515849 Syria
24: 9282173 Israel
25: 9282173 Israel
26: 7515849 Syria
27: 7515849 Syria
28: 9282173 Israel
29: 9282173 Israel
30: 1256116 David_Levy_(Israeli_politician)
31: 2820017 Kol_Yisrael
32: 9282173 Israel
33: 56472 Benjamin_Netanyahu
34: 7515865 Golan_Heights
35: 9282173 Israel
36: 7515849 Syria
37: 19323 Middle_East
38: 7515865 Golan_Heights
39: 7515865 Golan_Heights
40: 7515849 Syria
41: 7515849 Syria
42: 9282173 Israel
43: 5042765 God
44: 9282173 Israel
45: 3550889 Channel_2_(Israel)
46: 7515849 Syria
47: 9282173 Israel
48: 56472 Benjamin_Netanyahu
49: 56472 Benjamin_Netanyahu

50: 7515849 Syria
51: 3434750 United_States
1: 3434750 United_States
[winner: 108956 Washington,_D.C., pem_pos=0,
final_scores=-0.1608832, final_local_scores=-0.18168971,
pem=-1.7496481, similarity=8.182588]
RECALL MISS: 3434750 United_States]
2: 70843 Hafez_al-Assad
[winner: 364813 Bashar_al-Assad, pem_pos=0,
final_scores=0.14144321, final_local_scores=0.11669421,
pem=-0.48999527, similarity=10.236772]
gt: 70843 Hafez_al-Assad, pem_pos=1,
final_scores=0.10524599, final_local_scores=0.08885886,
pem=-1.3236102, similarity=13.939437]
3: 25391 Russia
[winner: 19004 Moscow, pem_pos=0,
final_scores=-0.38252425, final_local_scores=-0.42539078,
pem=-0.106582455, similarity=-8.025184]
RECALL MISS: 25391 Russia]
1: 9282173 Israel
[pem_pos=0, final_scores=0.5912915,
final_local_scores=0.53719264, pem=-0.21279564,
similarity=21.27783]
2: 7515849 Syria
[pem_pos=0, final_scores=0.6608917,
final_local_scores=0.6307466, pem=-0.26162508,
similarity=24.335209]
3: 16043 Jerusalem
[pem_pos=0, final_scores=0.49574065,
final_local_scores=0.4705106, pem=-0.13239029,
similarity=18.850647]
4: 9282173 Israel
[pem_pos=0, final_scores=0.58745295,
final_local_scores=0.5329996, pem=-0.21279564,
similarity=21.1527]
5: 7515849 Syria
[pem_pos=0, final_scores=0.7172323,
final_local_scores=0.6922901, pem=-0.26162508,
similarity=26.17179]
6: 17944843 Itamar_Rabinovich
[pem_pos=0, final_scores=0.529099,
final_local_scores=0.61552876, pem=0.0, similarity=22.458319]
7: 9282173 Israel
[pem_pos=0, final_scores=0.625261,
final_local_scores=0.57429916, pem=-0.21279564,
similarity=22.38516]
8: 108956 Washington,_D.C.
[pem_pos=0, final_scores=0.45317537,
final_local_scores=0.49357295, pem=-1.7496481,
similarity=28.333773]
9: 7515849 Syria
[pem_pos=0, final_scores=0.6605368,

```
final_local_scores=0.6303589, pem=-0.26162508,
similarity=24.323637]
10: 2820017 Kol_Yisrael
    [pem_pos=0, final_scores=0.37598962,
final_local_scores=0.41007423, pem=0.0, similarity=16.327147]
11: 7515849 Syria
    [pem_pos=19, final_scores=0.45889765,
final_local_scores=0.4100988, pem=-3.177024,
similarity=33.60503]
12: 7515849 Syria
    [pem_pos=0, final_scores=0.62298304,
final_local_scores=0.5893371, pem=-0.60890055,
similarity=24.988005]
13: 7515849 Syria
    [pem_pos=0, final_scores=0.65311044,
final_local_scores=0.62224674, pem=-0.32786387,
similarity=24.44177]
14: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.45177275,
final_local_scores=0.49204075, pem=-1.7496481,
similarity=28.288048]
15: 17944843 Itamar_Rabinovich
    [pem_pos=0, final_scores=0.91035455,
final_local_scores=1.0377548, pem=0.0, similarity=35.058384]
16: 8320263 Eliyahu_Ben-Elissar
    [pem_pos=0, final_scores=0.44930202,
final_local_scores=0.5467749, pem=0.0, similarity=20.406567]
17: 9282173 Israel
    [pem_pos=1, final_scores=0.60397965,
final_local_scores=0.5510525, pem=-0.34790626,
similarity=22.426186]
18: 8087628 Egypt
    [pem_pos=0, final_scores=0.5631081,
final_local_scores=0.53393537, pem=-0.27480945,
similarity=21.517866]
19: 206984 Likud
    [pem_pos=0, final_scores=0.45012742,
final_local_scores=0.47367662, pem=0.0, similarity=18.225168]
20: 9282173 Israel
    [pem_pos=0, final_scores=0.56743795,
final_local_scores=0.51113623, pem=-0.21279564,
similarity=20.500252]
21: 7515849 Syria
    [pem_pos=0, final_scores=0.6740386,
final_local_scores=0.6451075, pem=-0.26162508,
similarity=24.763765]
22: 7515849 Syria
    [pem_pos=19, final_scores=0.49917918,
final_local_scores=0.45410025, pem=-3.177024,
similarity=34.91812]
23: 7515849 Syria
    [pem_pos=0, final_scores=0.6193393,
```

```
final_local_scores=0.5853569, pem=-0.26162508,
similarity=22.980688]
24: 9282173 Israel
    [pem_pos=0, final_scores=0.62683076,
final_local_scores=0.57601386, pem=-0.21279564,
similarity=22.436333]
25: 9282173 Israel
    [pem_pos=1, final_scores=0.54027843,
final_local_scores=0.48146862, pem=-0.34790626,
similarity=20.349667]
26: 7515849 Syria
    [pem_pos=19, final_scores=0.48847377,
final_local_scores=0.44240618, pem=-3.177024,
similarity=34.569145]
27: 7515849 Syria
    [pem_pos=0, final_scores=0.5565855,
final_local_scores=0.5168079, pem=-0.60890055,
similarity=22.823587]
28: 9282173 Israel
    [pem_pos=0, final_scores=0.68164104,
final_local_scores=0.6358858, pem=-0.21279564,
similarity=24.22303]
29: 9282173 Israel
    [pem_pos=1, final_scores=0.6254265,
final_local_scores=0.5744799, pem=-0.34790626,
similarity=23.125307]
30: 1256116 David_Levy_(Israeli_politician)
    [pem_pos=1, final_scores=0.48010755,
final_local_scores=0.5230407, pem=-1.1434275,
similarity=25.916428]
31: 2820017 Kol_Yisrael
    [pem_pos=0, final_scores=0.38263458,
final_local_scores=0.41733283, pem=0.0, similarity=16.543756]
32: 9282173 Israel
    [pem_pos=0, final_scores=0.59656113,
final_local_scores=0.5429489, pem=-0.34790626,
similarity=22.184359]
33: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.6285799,
final_local_scores=0.6399206, pem=0.0, similarity=23.18622]
34: 7515865 Golan_Heights
    [pem_pos=0, final_scores=0.56902426,
final_local_scores=0.5657474, pem=0.0, similarity=20.972744]
35: 9282173 Israel
    [pem_pos=0, final_scores=0.65333647,
final_local_scores=0.6049673, pem=-0.21279564,
similarity=23.300358]
36: 7515849 Syria
    [pem_pos=0, final_scores=0.6297055,
final_local_scores=0.59668034, pem=-0.26162508,
similarity=23.318604]
37: 19323 Middle_East
```

[pem_pos=0, final_scores=0.46869993,
final_local_scores=0.45231736, pem=0.0, similarity=17.587765]
38: 7515865 Golan_Heights
[pem_pos=1, final_scores=0.3228181,
final_local_scores=0.2968046, pem=-1.3369088,
similarity=20.217272]
39: 7515865 Golan_Heights
[pem_pos=1, final_scores=0.3576452,
final_local_scores=0.33484793, pem=-1.3369088,
similarity=21.35256]
40: 7515849 Syria
[pem_pos=19, final_scores=0.39205992,
final_local_scores=0.3370887, pem=-3.177024,
similarity=31.426262]
41: 7515849 Syria
[pem_pos=0, final_scores=0.6180561,
final_local_scores=0.58395517, pem=-0.26162508,
similarity=22.938858]
42: 9282173 Israel
[pem_pos=0, final_scores=0.6146751,
final_local_scores=0.5627356, pem=-0.21279564,
similarity=22.040081]
43: 5042765 God
[pem_pos=0, final_scores=0.42181045,
final_local_scores=0.5123177, pem=-0.18657415,
similarity=20.392916]
44: 9282173 Israel
[pem_pos=0, final_scores=0.5880419,
final_local_scores=0.53364295, pem=-0.21279564,
similarity=21.171898]
45: 3550889 Channel_2_(Israel)
[pem_pos=1, final_scores=0.30797285,
final_local_scores=0.36840457, pem=-0.5544638,
similarity=18.0989]
46: 7515849 Syria
[pem_pos=19, final_scores=0.45658278,
final_local_scores=0.40757012, pem=-3.177024,
similarity=33.529568]
47: 9282173 Israel
[pem_pos=0, final_scores=0.5949362,
final_local_scores=0.5411739, pem=-0.21279564,
similarity=21.396637]
48: 56472 Benjamin_Netanyahu
[pem_pos=0, final_scores=0.48773438,
final_local_scores=0.48931164, pem=0.0, similarity=18.691748]
49: 56472 Benjamin_Netanyahu
[pem_pos=0, final_scores=0.4836275,
final_local_scores=0.4848255, pem=0.0, similarity=18.557873]
50: 7515849 Syria
[pem_pos=0, final_scores=0.6743117,
final_local_scores=0.6454059, pem=-0.26162508,
similarity=24.77267]

```
51: 3434750 United_States
    [pem_pos=0, final_scores=0.47892064,
final_local_scores=0.49541533, pem=-0.073457584,
similarity=19.273369]
1: 364813 Bashar_al-Assad
    [pem_pos=0, final_scores=0.14144321,
final_local_scores=0.11669421, pem=-0.48999527,
similarity=10.236772]

Sample: b'11_Polish'
[1 1 Polish 1] diplomat denies nurses stranded in [2 2 Libya
2] . [3 3 TUNIS 3] 1996-08-22 A [4 4 Polish 4] diplomat on
Thursday denied a [5 5 Polish 5] tabloid report this week
that [6 6 Libya 6] was refusing exit visas to 100 [7 7 Polish
7] nurses trying to return home after working in the North
African country . " This is not true . Up to today , we have
no knowledge of any nurse stranded or kept in [8 8 Libya 8]
without her will , and we have not received any complaint , "
the [9 9 Polish 9] embassy 's charge d'affaires in Tripoli ,
Tadeusz Awdankiewicz , told [10 10 Reuters 10] by telephone .
[11 11 Poland 11] 's labour ministry said this week it would
send a team to [12 12 Libya 12] to investigate , but
Awdankiewicz said the probe was prompted by some nurses
complaining about their work conditions such as non-payment
of their salaries . He said that there are an estimated 800 [1
3 13 Polish 13] nurses working in [14 14 Libya 14] .
1: 22936 Poland
2: 17633 Libya
3: 57659 Tunis
4: 22936 Poland
5: 22936 Poland
6: 17633 Libya
7: 22936 Poland
8: 17633 Libya
9: 22936 Poland
10: 18998750 Reuters
11: 22936 Poland
12: 17633 Libya
13: 22936 Poland
14: 17633 Libya
1: 22936 Poland
    [pem_pos=0, final_scores=0.63530236,
final_local_scores=0.6026254, pem=-0.81518245,
similarity=26.506346]
2: 17633 Libya
    [pem_pos=0, final_scores=0.454844,
final_local_scores=0.4228776, pem=-0.3054436,
similarity=18.370274]
3: 57659 Tunis
    [pem_pos=0, final_scores=0.411686,
final_local_scores=0.40485942, pem=-0.032719024,
similarity=16.349457]
```

```
4: 22936 Poland
    [pem_pos=0, final_scores=0.5993554,
final_local_scores=0.56335884, pem=-0.81518245,
similarity=25.334555]
5: 22936 Poland
    [pem_pos=0, final_scores=0.62749106,
final_local_scores=0.5940928, pem=-0.81518245,
similarity=26.251717]
6: 17633 Libya
    [pem_pos=0, final_scores=0.55235595,
final_local_scores=0.5293946, pem=-0.3054436,
similarity=21.548954]
7: 22936 Poland
    [pem_pos=0, final_scores=0.65498906,
final_local_scores=0.6241302, pem=-0.81518245,
similarity=27.148092]
8: 17633 Libya
    [pem_pos=0, final_scores=0.4886276,
final_local_scores=0.4597811, pem=-0.3054436,
similarity=19.471548]
9: 22936 Poland
    [pem_pos=0, final_scores=0.6661399,
final_local_scores=0.63631076, pem=-0.81518245,
similarity=27.511585]
10: 18998750 Reuters
    [pem_pos=0, final_scores=0.39679658,
final_local_scores=0.45614445, pem=0.0, similarity=17.701973]
11: 22936 Poland
    [pem_pos=0, final_scores=0.49680513,
final_local_scores=0.45133823, pem=-0.15333635,
similarity=18.392412]
12: 17633 Libya
    [pem_pos=0, final_scores=0.48969775,
final_local_scores=0.46095008, pem=-0.3054436,
similarity=19.506433]
13: 22936 Poland
    [pem_pos=0, final_scores=0.62756777,
final_local_scores=0.5941766, pem=-0.81518245,
similarity=26.254217]
14: 17633 Libya
    [pem_pos=0, final_scores=0.47565466,
final_local_scores=0.4456101, pem=-0.3054436,
similarity=19.048658]
```

Sample: b'12_Two'

Two [1 1 Iranian 1] opposition leaders meet in [2 2 Baghdad 2] . Hassan Hafidh [3 3 BAGHDAD 3] 1996-08-22 An [4 4 Iranian 4] exile group based in [5 5 Iraq 5] vowed on Thursday to extend support to [6 6 Iran 6] 's [7 7 Kurdish 7] rebels after they were attacked by [8 8 Iranian 8] troops deep inside [9 9 Iraq 9] last month . A Mujahideen Khalq statement said its leader [10 10 Massoud Rajavi 10] met in [11 11 Bagh

ad 11] the Secretary-General of the [12 12 Kurdistan Democrat ic Party of Iran 12] ([13 13 KDPI 13]) Hassan Rastegar on Wednesday and voiced his support to [14 14 Iran 14] 's rebel [15 15 Kurds 15] . " Rajavi emphasised that the [16 16 Iranian 16] [17 17 Resistance 17] would continue to stand side by side with their Kurdish compatriots and the resistance movement in [18 18 Iranian Kurdistan 18] , " it said . A spokesman for the group said the meeting " signals a new level of cooperation between Mujahideen Khalq and the [19 19 I ranian Kurdish 19] oppositions " . [20 20 Iran 20] heavily bombarded targets in northern [21 21 Iraq 21] in July in pursuit of [22 22 KDPI 22] guerrillas based in [23 23 Iraqi Ku rdish 23] areas outside the control of the government in [24 2 4 Baghdad 24] . [25 25 Iraqi Kurdish 25] areas bordering [26 2 6 Iran 26] are under the control of guerrillas of the Iraqi Kurdish Patriotic Union of Kurdistan ([27 27 PUK 27]) group . [28 28 PUK 28] and [29 29 Iraq 29] 's [30 30 Kurdistan Democratic Party 30] ([31 31 KDP 31]) the two main [32 32 Ir aqi Kurdish 32] factions , have had northern [33 33 Iraq 33] under their control since [34 34 Iraqi 34] forces were ousted from [35 35 Kuwait 35] in the 1991 [36 36 Gulf War 36] . Clashes between the two parties broke out at the weekend in the most serious fighting since a U.S.-sponsored ceasefire last year . Mujahideen Khalq said [37 37 Iranian 37] troops had also been shelling [38 38 KDP 38] positions in [39 39 Qas ri 39] region in [40 40 Suleimaniya 40] province near the [41 41 Iranian 41] border over the last two days . It said about 100 [42 42 Iraqi Kurds 42] were killed or wounded in the attack . Both [43 43 Iran 43] and Turkey mount air and land strikes at targets in northern [44 44 Iraq 44] in pursuit of their own [45 45 Kurdish 45] rebels . A U.S.-led air force in southern Turkey protects [46 46 Iraqi Kurds 46] from possible attacks by [47 47 Baghdad 47] troops .

1: 14653 Iran
2: 4492 Baghdad
3: 4492 Baghdad
4: 14653 Iran
5: 7515928 Iraq
6: 14653 Iran
7: 17068 Kurdish_people
8: 14653 Iran
9: 7515928 Iraq
10: 5557401 Massoud_Rajavi
11: 4492 Baghdad
12: 3197182 Democratic_Party_of_Iranian_Kurdistan
13: 3197182 Democratic_Party_of_Iranian_Kurdistan
14: 14653 Iran
15: 17068 Kurdish_people
16: 14653 Iran
17: 83933 Resistance_movement
18: 24993749 Iranian_Kurdistan
19: 24993749 Iranian_Kurdistan

```
20: 14653 Iran
21: 7515928 Iraq
22: 3197182 Democratic_Party_of_Iranian_Kurdistan
23: 679693 Iraqi_Kurdistan
24: 4492 Baghdad
25: 679693 Iraqi_Kurdistan
26: 14653 Iran
27: 70789 Patriotic_Union_of_Kurdistan
28: 70789 Patriotic_Union_of_Kurdistan
29: 7515928 Iraq
30: 25659771 Kurdistan_Democratic_Party
31: 25659771 Kurdistan_Democratic_Party
32: 679693 Iraqi_Kurdistan
33: 7515928 Iraq
34: 7515928 Iraq
35: 7515890 Kuwait
36: 182000 Gulf_War
37: 14653 Iran
38: 25659771 Kurdistan_Democratic_Party
39: 5544058 Qasr-e Shirin
40: 653537 Sulaymaniyah
41: 14653 Iran
42: 679693 Iraqi_Kurdistan
43: 14653 Iran
44: 7515928 Iraq
45: 17068 Kurdish_people
46: 679693 Iraqi_Kurdistan
47: 7515928 Iraq
1: 14653 Iran
    [pem_pos=0, final_scores=0.5405546,
final_local_scores=0.5468292, pem=-0.57066226,
similarity=23.511538]
2: 4492 Baghdad
    [pem_pos=0, final_scores=0.4170258,
final_local_scores=0.37223017, pem=0.0, similarity=15.197804]
3: 4492 Baghdad
    [pem_pos=0, final_scores=0.45537525,
final_local_scores=0.4141211, pem=0.0, similarity=16.447912]
4: 14653 Iran
    [pem_pos=0, final_scores=0.49071175,
final_local_scores=0.49238348, pem=-0.57066226,
similarity=21.886768]
5: 7515928 Iraq
    [pem_pos=0, final_scores=0.58945805,
final_local_scores=0.54580957, pem=-0.25379938,
similarity=21.75796]
6: 14653 Iran
    [pem_pos=0, final_scores=0.6436185,
final_local_scores=0.65941095, pem=-0.13856095,
similarity=24.521368]
7: 17068 Kurdish_people
    [pem_pos=1, final_scores=0.2179328,
```

```
final_local_scores=0.157285, pem=-0.713598,
similarity=12.664068]
8: 14653 Iran
    [pem_pos=0, final_scores=0.5127079,
final_local_scores=0.51641095, pem=-0.57066226,
similarity=22.603796]
9: 7515928 Iraq
    [pem_pos=0, final_scores=0.60316926,
final_local_scores=0.56078696, pem=-0.25379938,
similarity=22.204916]
10: 5557401 Massoud_Rajavi
    [pem_pos=0, final_scores=0.5076584,
final_local_scores=0.57927495, pem=0.0, similarity=21.376432]
11: 4492 Baghdad
    [pem_pos=0, final_scores=0.3840044,
final_local_scores=0.33615932, pem=0.0, similarity=14.121378]
12: 3197182 Democratic_Party_of_Iranian_Kurdistan
    [pem_pos=0, final_scores=0.4368524,
final_local_scores=0.45721257, pem=0.0, similarity=17.733849]
13: 3197182 Democratic_Party_of_Iranian_Kurdistan
    [pem_pos=0, final_scores=0.5780985,
final_local_scores=0.6115024, pem=0.0, similarity=22.338165]
14: 14653 Iran
    [pem_pos=0, final_scores=0.69126207,
final_local_scores=0.7114542, pem=-0.13856095,
similarity=26.074442]
15: 17068 Kurdish_people
    [pem_pos=0, final_scores=0.23619173,
final_local_scores=0.17723009, pem=0.0, similarity=9.3786125]
16: 14653 Iran
    [pem_pos=0, final_scores=0.5288409,
final_local_scores=0.5340338, pem=-0.57066226,
similarity=23.129696]
17: 83933 Resistance_movement
    [pem_pos=8, final_scores=0.12177919,
final_local_scores=0.14838201, pem=-2.9856632,
similarity=24.754229]
18: 24993749 Iranian_Kurdistan
    [pem_pos=0, final_scores=0.37508643,
final_local_scores=0.34877726, pem=0.0, similarity=14.497922]
19: 24993749 Iranian_Kurdistan
    [pem_pos=0, final_scores=0.45941025,
final_local_scores=0.44088823, pem=0.0, similarity=17.246696]
20: 14653 Iran
    [pem_pos=0, final_scores=0.60382706,
final_local_scores=0.61594448, pem=-0.13856095,
similarity=23.22425]
21: 7515928 Iraq
    [pem_pos=0, final_scores=0.6024864,
final_local_scores=0.560041, pem=-0.25379938,
similarity=22.182655]
22: 3197182 Democratic_Party_of_Iranian_Kurdistan
```

[pem_pos=0, final_scores=0.5883498,
final_local_scores=0.62270045, pem=0.0, similarity=22.672337]
23: 679693 Iraqi_Kurdistan
[pem_pos=0, final_scores=0.46667916,
final_local_scores=0.4246878, pem=0.0, similarity=16.763245]
24: 4492 Baghdad
[pem_pos=0, final_scores=0.38639933,
final_local_scores=0.33877543, pem=0.0, similarity=14.199448]
25: 679693 Iraqi_Kurdistan
[pem_pos=0, final_scores=0.4354841,
final_local_scores=0.39061195, pem=0.0, similarity=15.746352]
26: 14653 Iran
[pem_pos=0, final_scores=0.5800291,
final_local_scores=0.5899491, pem=-0.13856095,
similarity=22.448486]
27: 70789 Patriotic_Union_of_Kurdistan
[pem_pos=0, final_scores=0.5658477,
final_local_scores=0.586544, pem=0.0, similarity=21.593355]
28: 70789 Patriotic_Union_of_Kurdistan
[pem_pos=0, final_scores=0.53626364,
final_local_scores=0.5542279, pem=0.0, similarity=20.628979]
29: 7515928 Iraq
[pem_pos=0, final_scores=0.619762,
final_local_scores=0.578912, pem=-0.25379938,
similarity=22.745804]
30: 25659771 Kurdistan_Democratic_Party
[pem_pos=0, final_scores=0.59479564,
final_local_scores=0.60535157, pem=0.0, similarity=22.154612]
31: 25659771 Kurdistan_Democratic_Party
[pem_pos=0, final_scores=0.5866565,
final_local_scores=0.5964608, pem=0.0, similarity=21.889294]
32: 679693 Iraqi_Kurdistan
[pem_pos=0, final_scores=0.45158976,
final_local_scores=0.40820497, pem=0.0, similarity=16.271364]
33: 7515928 Iraq
[pem_pos=0, final_scores=0.5658825,
final_local_scores=0.5200569, pem=-0.25379938,
similarity=20.989449]
34: 7515928 Iraq
[pem_pos=0, final_scores=0.59244716,
final_local_scores=0.5490747, pem=-0.48469657,
similarity=23.111053]
35: 7515890 Kuwait
[pem_pos=0, final_scores=0.4998455,
final_local_scores=0.48741663, pem=-0.3372397,
similarity=20.469158]
36: 182000 Gulf_War
[pem_pos=0, final_scores=0.36374134,
final_local_scores=0.37790108, pem=0.0, similarity=15.367034]
37: 14653 Iran
[pem_pos=0, final_scores=0.6246663,
final_local_scores=0.6387084, pem=-0.57066226,

```
similarity=26.253395]
38: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.6692854,
final_local_scores=0.6867203, pem=0.0, similarity=24.582817]
39: 5544058 Qasr-e Shirin
    [pem_pos=0, final_scores=0.5406924,
final_local_scores=0.5915013, pem=0.0, similarity=21.741291]
40: 653537 Sulaymaniyah
    [pem_pos=0, final_scores=0.32384044,
final_local_scores=0.30082366, pem=0.0, similarity=13.066891]
41: 14653 Iran
    [pem_pos=0, final_scores=0.5128892,
final_local_scores=0.51660895, pem=-0.57066226,
similarity=22.609705]
42: 679693 Iraqi_Kurdistan
    [pem_pos=0, final_scores=0.34045708,
final_local_scores=0.28680936, pem=0.0, similarity=12.648676]
43: 14653 Iran
    [pem_pos=0, final_scores=0.59988946,
final_local_scores=0.61164355, pem=-0.13856095,
similarity=23.095892]
44: 7515928 Iraq
    [pem_pos=0, final_scores=0.6064721,
final_local_scores=0.56439483, pem=-0.25379938,
similarity=22.312582]
45: 17068 Kurdish_people
    [pem_pos=1, final_scores=0.19162975,
final_local_scores=0.12855291, pem=-0.713598,
similarity=11.806645]
46: 679693 Iraqi_Kurdistan
    [pem_pos=0, final_scores=0.28671712,
final_local_scores=0.22810665, pem=0.0, similarity=10.896871]
47: 7515928 Iraq
    [pem_pos=8, final_scores=0.29444987,
final_local_scores=0.22355795, pem=-2.5644467,
similarity=24.706987]
```

Sample: b'13_Saudi'

[1 1 Saudi 1] riyal rates steady in quiet summer trade . [2 2
MANAMA 2] 1996-08-22 The spot [3 3 Saudi 3] riyal against the
dollar and riyal interbank deposit rates were mainly steady
this week in quiet summer trade , dealers in the kingdom
said . " There were no changes in [4 4 Saudi 4] riyal rates .
The market was very quiet because of summer holidays , " one
dealer said . The spot riyal was put at 3.7504 / 06 to the
dollar . One-month interbank deposits were at 5-1/2 , 3/8
percent , three months were 5-5/8 , 1/2 percent and six
months were 5-3/4 , 5/8 percent . One-year funds were at
six , 5-7/8 percent .

1: 349303 Saudi_Arabia

2: 20481 Manama

3: 349303 Saudi_Arabia

```
4: 349303 Saudi_Arabia
1: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.54183334,
final_local_scores=0.47461134, pem=0.0, similarity=18.253061]
2: 20481 Manama
    [pem_pos=0, final_scores=0.41678375,
final_local_scores=0.42483622, pem=0.0, similarity=16.767673]
3: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.56734717,
final_local_scores=0.50248134, pem=0.0, similarity=19.084759]
4: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.5923751,
final_local_scores=0.52982056, pem=0.0, similarity=19.900616]
```

Sample: b'14_Israel'

[1 1 Israel 1] approves [2 2 Arafat 2] 's flight to West Bank . [3 3 JERUSALEM 3] 1996-08-22 [4 4 Israel 4] gave [5 5 Palestinian 5] President [6 6 Yasser Arafat 6] permission on Thursday to fly over its territory to the West Bank , ending a brief Israeli-PLO crisis , an [7 7 Arafat 7] adviser said . " The problem is over . The president 's aircraft has received permission to pass through [8 8 Israeli 8] airspace but the president is not expected to travel to the West Bank before Monday , " Nabil Abu Rdainah told [9 9 Reuters 9] . [10 10 Arafat 10] had been scheduled to meet former [11 11 Israeli 11] prime minister [12 12 Shimon Peres 12] in the West Bank town of [13 13 Ramallah 13] on Thursday but the venue was changed to [14 14 Gaza 14] after [15 15 Israel 15] denied flight clearance to the [16 16 Palestinian 16] leader 's helicopters . [17 17 Palestinian 17] officials accused right-wing Prime Minister [18 18 Benjamin Netanyahu 18] of trying to stop the [19 19 Ramallah 19] meeting by keeping [20 20 Arafat 20] grounded . [21 21 Arafat 21] subsequently cancelled a meeting between [22 22 Israeli 22] and PLO officials , on civilian affairs , at the [23 23 Allenby Bridge 23] crossing between [1 1 Jordan 1] and the West Bank . Abu Rdainah said [2 4 24 Arafat 24] had decided against flying to the West Bank on Thursday , after [25 25 Israel 25] lifted the ban , because he had a busy schedule in [26 26 Gaza 26] and would not be free until Monday .

```
1: 9282173 Israel
2: 34289 Yasser_Arafat
3: 16043 Jerusalem
4: 9282173 Israel
5: 241405 State_of_Palestine
6: 34289 Yasser_Arafat
7: 34289 Yasser_Arafat
8: 9282173 Israel
9: 18998750 Reuters
10: 34289 Yasser_Arafat
11: 9282173 Israel
12: 48569 Shimon_Peres
```

13: 44132 Ramallah
14: 12047 Gaza_Strip
15: 9282173 Israel
16: 241405 State_of_Palestine
17: 241405 State_of_Palestine
18: 56472 Benjamin_Netanyahu
19: 44132 Ramallah
20: 34289 Yasser_Arafat
21: 34289 Yasser_Arafat
22: 9282173 Israel
23: 4847484 Allenby_Bridge
24: 34289 Yasser_Arafat
25: 9282173 Israel
26: 12047 Gaza_Strip
1: 47910 Jordan_River
[winner: 7515964 Jordan, pem_pos=0,
final_scores=0.25156933, final_local_scores=0.20830771,
pem=-0.53647494, similarity=13.223465]
gt: 47910 Jordan_River, pem_pos=5,
final_scores=0.071642175, final_local_scores=0.08081025,
pem=-4.233828, similarity=29.525467]
1: 9282173 Israel
[pem_pos=0, final_scores=0.52251405,
final_local_scores=0.4856195, pem=-0.21279564,
similarity=19.738783]
2: 34289 Yasser_Arafat
[pem_pos=0, final_scores=0.64341253,
final_local_scores=0.59994423, pem=0.0, similarity=21.993246]
3: 16043 Jerusalem
[pem_pos=0, final_scores=0.51501596,
final_local_scores=0.5028905, pem=-0.13239029,
similarity=19.816929]
4: 9282173 Israel
[pem_pos=0, final_scores=0.5779844,
final_local_scores=0.5462124, pem=-0.21279564,
similarity=21.546997]
5: 241405 State_of_Palestine
[pem_pos=10, final_scores=0.6424306,
final_local_scores=0.6442199, pem=-2.8508017,
similarity=38.817623]
6: 34289 Yasser_Arafat
[pem_pos=0, final_scores=0.6369137,
final_local_scores=0.5911954, pem=0.0, similarity=21.732162]
7: 34289 Yasser_Arafat
[pem_pos=0, final_scores=0.5877041,
final_local_scores=0.5390912, pem=0.0, similarity=20.17727]
8: 9282173 Israel
[pem_pos=0, final_scores=0.44508713,
final_local_scores=0.40104228, pem=-0.34790626,
similarity=17.949583]
9: 18998750 Reuters
[pem_pos=0, final_scores=0.39927047,

```
final_local_scores=0.41887015, pem=0.0, similarity=16.589634]
10: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.64451545,
final_local_scores=0.60114896, pem=0.0, similarity=22.029198]
11: 9282173 Israel
    [pem_pos=0, final_scores=0.59650767,
final_local_scores=0.56644624, pem=-0.34790626,
similarity=22.885569]
12: 48569 Shimon_Peres
    [pem_pos=0, final_scores=0.5830575,
final_local_scores=0.5755448, pem=0.0, similarity=21.265118]
13: 44132 Ramallah
    [pem_pos=0, final_scores=0.50710946,
final_local_scores=0.47748828, pem=0.0, similarity=18.338915]
14: 12047 Gaza_Strip
    [pem_pos=1, final_scores=0.40244663,
final_local_scores=0.3655032, pem=-1.374985,
similarity=22.47444]
15: 9282173 Israel
    [pem_pos=0, final_scores=0.6217895,
final_local_scores=0.59406286, pem=-0.21279564,
similarity=22.97495]
16: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.71022147,
final_local_scores=0.71827114, pem=-2.8508017,
similarity=41.02746]
17: 241405 State_of_Palestine
    [pem_pos=10, final_scores=0.6289257,
final_local_scores=0.62946784, pem=-2.8508017,
similarity=38.37739]
18: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.6333907,
final_local_scores=0.6243462, pem=0.0, similarity=22.721449]
19: 44132 Ramallah
    [pem_pos=0, final_scores=0.4057985,
final_local_scores=0.36682135, pem=0.0, similarity=15.036393]
20: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.6479335,
final_local_scores=0.60488266, pem=0.0, similarity=22.140617]
21: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.6421936,
final_local_scores=0.5986127, pem=0.0, similarity=21.95351]
22: 9282173 Israel
    [pem_pos=0, final_scores=0.5463006,
final_local_scores=0.5116027, pem=-0.34790626,
similarity=21.248926]
23: 4847484 Allenby_Bridge
    [pem_pos=0, final_scores=0.43321985,
final_local_scores=0.49947095, pem=0.0, similarity=18.994923]
24: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.64397943,
final_local_scores=0.60056347, pem=0.0, similarity=22.011724]
```

```
25: 9282173 Israel
    [pem_pos=0, final_scores=0.64142615,
final_local_scores=0.6155129, pem=-0.21279564,
similarity=23.61506]
26: 12047 Gaza_Strip
    [pem_pos=1, final_scores=0.33020818,
final_local_scores=0.28659368, pem=-1.374985,
similarity=20.119621]
1: 7515964 Jordan
    [pem_pos=0, final_scores=0.25156933,
final_local_scores=0.20830771, pem=-0.53647494,
similarity=13.223465]
```

Sample: b'15_Arafat'

[1 1 Arafat 1] to meet [2 2 Peres 2] in [3 3 Gaza 3] after flight ban . [4 4 JERUSALEM 4] 1996-08-22 [5 5 Yasser Arafat 5] will meet [6 6 Shimon Peres 6] in [7 7 Gaza 7] on Thursday after [8 8 Palestinians 8] said the right-wing [9 9 Israeli 9] government had barred the [10 10 Palestinian 10] leader from flying to the West Bank for talks with the former prime minister . " The meeting between [11 11 Peres 11] and [12 12 Arafat 12] will take place at [13 13 Erez 13] checkpoint in [14 14 Gaza 14] and not in [15 15 Ramallah 15] as planned , " [16 16 Peres 16] ' office said . [17 17 Palestinian 17] officials said the [18 18 Israeli 18] government had barred [19 19 Arafat 19] from overflying [20 20 Israel 20] in a [21 21 Palestinian 21] helicopter to the West Bank in an attempt to bar the meeting with [22 22 Peres 22] . [23 23 Israeli 23] Prime Minister [24 24 Benjamin Netanyahu 24] has accused opposition leader [25 25 Peres 25] , who he defeated in May elections , of trying to undermine his [26 26 Likud 26] government 's authority to conduct peace talks .

```
1: 34289 Yasser_Arafat
2: 48569 Shimon_Peres
3: 12047 Gaza_Strip
4: 16043 Jerusalem
5: 34289 Yasser_Arafat
6: 48569 Shimon_Peres
7: 12047 Gaza_Strip
8: 241405 State_of_Palestine
9: 9282173 Israel
10: 241405 State_of_Palestine
11: 48569 Shimon_Peres
12: 34289 Yasser_Arafat
13: 14895674 Erez
14: 12047 Gaza_Strip
15: 44132 Ramallah
16: 48569 Shimon_Peres
17: 241405 State_of_Palestine
18: 9282173 Israel
19: 34289 Yasser_Arafat
20: 9282173 Israel
```

```
21: 241405 State_of_Palestine
22: 48569 Shimon_Peres
23: 9282173 Israel
24: 56472 Benjamin_Netanyahu
25: 48569 Shimon_Peres
26: 206984 Likud
1: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.572498,
final_local_scores=0.52636534, pem=0.0, similarity=19.797504]
2: 48569 Shimon_Peres
    [pem_pos=0, final_scores=0.7876032,
final_local_scores=0.7655027, pem=0.0, similarity=26.93384]
3: 12047 Gaza_Strip
    [pem_pos=1, final_scores=0.31517553,
final_local_scores=0.28234917, pem=-1.374985,
similarity=19.992958]
4: 16043 Jerusalem
    [pem_pos=0, final_scores=0.517013,
final_local_scores=0.5230569, pem=-0.13239029,
similarity=20.418734]
5: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.56545264,
final_local_scores=0.51676786, pem=0.0, similarity=19.511097]
6: 48569 Shimon_Peres
    [pem_pos=0, final_scores=0.6077236,
final_local_scores=0.56684965, pem=0.0, similarity=21.005636]
7: 12047 Gaza_Strip
    [pem_pos=1, final_scores=0.32717866,
final_local_scores=0.29546076, pem=-1.374985,
similarity=20.384233]
8: 241405 State_of_Palestine
    [pem_pos=5, final_scores=0.68446714,
final_local_scores=0.6949742, pem=-1.7909594,
similarity=34.568645]
9: 9282173 Israel
    [pem_pos=1, final_scores=0.62727004,
final_local_scores=0.61367863, pem=-0.34790626,
similarity=24.295074]
10: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.6353591,
final_local_scores=0.6413311, pem=-2.8508017,
similarity=38.731415]
11: 48569 Shimon_Peres
    [pem_pos=0, final_scores=0.7353312,
final_local_scores=0.70840347, pem=0.0, similarity=25.229885]
12: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.62841296,
final_local_scores=0.58744395, pem=0.0, similarity=21.620213]
13: 14895674 Erez
    [pem_pos=0, final_scores=0.25230807,
final_local_scores=0.33295456, pem=-0.058208905,
similarity=14.34229]
```

```
14: 12047 Gaza_Strip
    [pem_pos=1, final_scores=0.32154202,
final_local_scores=0.2893036, pem=-1.374985,
similarity=20.20049]
15: 44132 Ramallah
    [pem_pos=0, final_scores=0.33244646,
final_local_scores=0.30621675, pem=0.0, similarity=13.227832]
16: 48569 Shimon_Peres
    [pem_pos=0, final_scores=0.77891546,
final_local_scores=0.7560126, pem=0.0, similarity=26.650637]
17: 241405 State_of_Palestine
    [pem_pos=10, final_scores=0.56470436,
final_local_scores=0.5641515, pem=-2.8508017,
similarity=36.428223]
18: 9282173 Israel
    [pem_pos=1, final_scores=0.59618264,
final_local_scores=0.5797203, pem=-0.34790626,
similarity=23.28169]
19: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.65481,
final_local_scores=0.6162787, pem=0.0, similarity=22.4807]
20: 9282173 Israel
    [pem_pos=0, final_scores=0.52955526,
final_local_scores=0.50694007, pem=-0.21279564,
similarity=20.37503]
21: 241405 State_of_Palestine
    [pem_pos=10, final_scores=0.5442129,
final_local_scores=0.5417677, pem=-2.8508017,
similarity=35.760246]
22: 48569 Shimon_Peres
    [pem_pos=0, final_scores=0.6866996,
final_local_scores=0.6552809, pem=0.0, similarity=23.644602]
23: 9282173 Israel
    [pem_pos=0, final_scores=0.5673564,
final_local_scores=0.548232, pem=-0.34790626,
similarity=22.342018]
24: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.7286894,
final_local_scores=0.70957166, pem=0.0, similarity=25.264746]
25: 48569 Shimon_Peres
    [pem_pos=0, final_scores=0.69428325,
final_local_scores=0.6635648, pem=0.0, similarity=23.89181]
26: 206984 Likud
    [pem_pos=0, final_scores=0.4870966,
final_local_scores=0.48754615, pem=0.0, similarity=18.639063]
```

Sample: b'16_Afghan'

[1 1 Afghan 1] [2 2 UAE 2] embassy says [3 3 Taliban 3]
guards going home . Hilary Gush [4 4 DUBAI 4] 1996-08-22
Three [5 5 Afghan 5] guards brought to the [6 6 United Arab Em
irates 6] last week by [7 7 Russian 7] hostages who escaped
from the [8 8 Taliban 8] militia will return to [9 9 Afghanis

tan 9] in a few days , the [10 10 Afghan 10] embassy in [11 11 Abu Dhabi 11] said on Thursday . " Our ambassador is in touch with the [12 12 UAE 12] foreign ministry . Their return to [13 13 Afghanistan 13] will take place in two or three days , " an embassy official said . " The embassy is issuing them travel documents for their return to their homeland . There is no objection to their travel , " he added . The three Islamic Taliban guards were overpowered by seven [14 14 Russian 14] aircrew who escaped to [15 15 UAE 15] state [16 16 Sharjah 16] last Friday on board their own aircraft after a year in the captivity of [17 17 Taliban 17] militia in [18 18 Kandahar 18] in southern [19 19 Afghanistan 19] . The [20 20 UAE 20] said on Monday it would hand over the three to the International Red Crescent , possibly last Tuesday . It has since been silent on the issue . When asked whether the three guards would travel back to [21 21 Kandahar 21] or the [22 22 Afghan 22] capital [23 23 Kabul 23] , the embassy official said : " That has not been decided , but possibly [24 24 Kandahar 24] . " [25 25 Kandahar 25] is the headquarters of the opposition [26 26 Taliban 26] militia . [27 27 Kabul 27] is controlled by President [28 28 Burhanuddin Rabbani 28] 's government , which [29 29 Taliban 29] is fighting to overthrow . The embassy official said the three men , believed to be in their 20s , were currently in [1 1 Abu Dhabi 1] . He did not elaborate . The [30 30 Russians 30] , working for the Aerostan firm in the [31 31 Russian 31] republic of [32 32 Tatarstan 32] , were taken hostage after a [33 33 Taliban 33] MiG-19 fighter forced their cargo plane to land in August 1995 . [34 34 Taliban 34] said its shipment of ammunition from [35 35 Albania 35] was evidence of [36 36 Russian 36] military support for [37 37 Rabbani 37] 's government . 2 Moscow said the crew 's nationality was coincidental . Numerous diplomatic attempts to free the seven failed . The [38 38 Russians 38] , who said they overpowered the guards -- two armed with Kalashnikov automatic rifles -- while doing regular maintenance work on their Ilyushin 76 cargo plane last Friday , left the [39 39 UAE 39] capital [40 40 Abu Dhabi 40] for home on Sunday .

1: 737 Afghanistan
2: 69328 United_Arab_Emirates
3: 30635 Taliban
4: 211583 Dubai
5: 737 Afghanistan
6: 69328 United_Arab_Emirates
7: 25391 Russia
8: 30635 Taliban
9: 737 Afghanistan
10: 737 Afghanistan
11: 18950756 Abu_Dhabi
12: 69328 United_Arab_Emirates
13: 737 Afghanistan
14: 25391 Russia

```
15: 69328 United_Arab_Emirates
16: 210589 Sharjah_(emirate)
17: 30635 Taliban
18: 17260 Kandahar
19: 737 Afghanistan
20: 69328 United_Arab_Emirates
21: 17260 Kandahar
22: 737 Afghanistan
23: 16826 Kabul
24: 17260 Kandahar
25: 17260 Kandahar
26: 30635 Taliban
27: 16826 Kabul
28: 4611 Burhanuddin_Rabbani
29: 30635 Taliban
30: 25391 Russia
31: 25391 Russia
32: 67135 Tatarstan
33: 30635 Taliban
34: 30635 Taliban
35: 738 Albania
36: 25391 Russia
37: 4611 Burhanuddin_Rabbani
38: 25391 Russia
39: 69328 United_Arab_Emirates
40: 18950756 Abu_Dhabi
1: 8821238 Abu_Dhabi_(emirate)
    [winner: 18950756 Abu_Dhabi, pem_pos=0,
final_scores=0.19111474, final_local_scores=0.17287874,
pem=-0.065649725, similarity=9.605773]
    gt: 8821238 Abu_Dhabi_(emirate), pem_pos=1,
final_scores=0.16081677, final_local_scores=0.1931062,
pem=-1.6919055, similarity=19.053232 ]
2: 25391 Russia
    [winner: 19004 Moscow, pem_pos=0,
final_scores=-0.24363627, final_local_scores=-0.28398162,
pem=-0.106582455, similarity=-3.8052535|
     RECALL MISS: 25391 Russia]
1: 737 Afghanistan
    [pem_pos=0, final_scores=0.56704867,
final_local_scores=0.51959795, pem=-0.869566,
similarity=24.324387]
2: 69328 United_Arab_Emirates
    [pem_pos=0, final_scores=0.51526326,
final_local_scores=0.48900658, pem=-0.1950423,
similarity=19.743315]
3: 30635 Taliban
    [pem_pos=0, final_scores=0.45095044,
final_local_scores=0.4307348, pem=0.0, similarity=16.943699]
4: 211583 Dubai
    [pem_pos=0, final_scores=0.49383783,
final_local_scores=0.50590384, pem=-0.08179128,
```

```
similarity=19.631687]
5: 737 Afghanistan
    [pem_pos=0, final_scores=0.59436446,
final_local_scores=0.5494363, pem=-0.869566,
similarity=25.214823]
6: 69328 United_Arab_Emirates
    [pem_pos=0, final_scores=0.37184376,
final_local_scores=0.33234254, pem=-0.24986836,
similarity=15.3663]
7: 25391 Russia
    [pem_pos=1, final_scores=0.7591829,
final_local_scores=0.7965082, pem=-1.0515022,
similarity=33.57734]
8: 30635 Taliban
    [pem_pos=0, final_scores=0.50828147,
final_local_scores=0.49336028, pem=0.0, similarity=18.812569]
9: 737 Afghanistan
    [pem_pos=0, final_scores=0.6167721,
final_local_scores=0.5739133, pem=-0.22286008,
similarity=22.428377]
10: 737 Afghanistan
    [pem_pos=0, final_scores=0.5871652,
final_local_scores=0.54157215, pem=-0.869566,
similarity=24.980143]
11: 18950756 Abu_Dhabi
    [pem_pos=0, final_scores=0.28448826,
final_local_scores=0.2722885, pem=-0.065649725,
similarity=12.572358]
12: 69328 United_Arab_Emirates
    [pem_pos=0, final_scores=0.6189078,
final_local_scores=0.6022225, pem=-0.1950423,
similarity=23.121902]
13: 737 Afghanistan
    [pem_pos=0, final_scores=0.63912505,
final_local_scores=0.5983304, pem=-0.22286008,
similarity=23.157032]
14: 25391 Russia
    [pem_pos=1, final_scores=0.77022225,
final_local_scores=0.80856705, pem=-1.0515022,
similarity=33.9372]
15: 69328 United_Arab_Emirates
    [pem_pos=0, final_scores=0.47366202,
final_local_scores=0.44356352, pem=-0.1950423,
similarity=18.387203]
16: 210589 Sharjah_(emirate)
    [pem_pos=1, final_scores=0.24693541,
final_local_scores=0.24872106, pem=-0.42679185,
similarity=13.833006]
17: 30635 Taliban
    [pem_pos=0, final_scores=0.46666658,
final_local_scores=0.44790232, pem=0.0, similarity=17.45601]
18: 17260 Kandahar
```

```
[pem_pos=0, final_scores=0.3258344,
final_local_scores=0.2910273, pem=-0.11192558,
similarity=13.383217]
19: 737 Afghanistan
    [pem_pos=0, final_scores=0.61049837,
final_local_scores=0.5670602, pem=-0.22286008,
similarity=22.223867]
20: 69328 United_Arab_Emirates
    [pem_pos=0, final_scores=0.5009064,
final_local_scores=0.47332388, pem=-0.1950423,
similarity=19.275312]
21: 17260 Kandahar
    [pem_pos=0, final_scores=0.35024798,
final_local_scores=0.31769544, pem=-0.11192558,
similarity=14.179048]
22: 737 Afghanistan
    [pem_pos=0, final_scores=0.5149495,
final_local_scores=0.46268755, pem=-0.869566,
similarity=22.626068]
23: 16826 Kabul
    [pem_pos=0, final_scores=0.43531448,
final_local_scores=0.39127916, pem=0.0, similarity=15.766263]
24: 17260 Kandahar
    [pem_pos=0, final_scores=0.33421308,
final_local_scores=0.3001797, pem=-0.11192558,
similarity=13.6563425]
25: 17260 Kandahar
    [pem_pos=0, final_scores=0.34932858,
final_local_scores=0.31669113, pem=-0.11192558,
similarity=14.149076]
26: 30635 Taliban
    [pem_pos=0, final_scores=0.4332041,
final_local_scores=0.4113497, pem=0.0, similarity=16.36521]
27: 16826 Kabul
    [pem_pos=0, final_scores=0.4121186,
final_local_scores=0.36594117, pem=0.0, similarity=15.010128]
28: 4611 Burhanuddin_Rabbani
    [pem_pos=0, final_scores=0.53158975,
final_local_scores=0.5605609, pem=0.0, similarity=20.817968]
29: 30635 Taliban
    [pem_pos=0, final_scores=0.47859603,
final_local_scores=0.46093345, pem=0.0, similarity=17.844887]
30: 25391 Russia
    [pem_pos=1, final_scores=0.6272785,
final_local_scores=0.6524227, pem=-1.8802736,
similarity=33.784527]
31: 25391 Russia
    [pem_pos=1, final_scores=0.67653924,
final_local_scores=0.70623267, pem=-1.0515022,
similarity=30.883335]
32: 67135 Tatarstan
    [pem_pos=0, final_scores=0.39515084,
```

```
final_local_scores=0.41943246, pem=0.0, similarity=16.606415]
33: 30635 Taliban
    [pem_pos=0, final_scores=0.42122346,
final_local_scores=0.39826262, pem=0.0, similarity=15.974664]
34: 30635 Taliban
    [pem_pos=0, final_scores=0.33195585,
final_local_scores=0.30075136, pem=0.0, similarity=13.0647335]
35: 738 Albania
    [pem_pos=0, final_scores=0.64136094,
final_local_scores=0.6721127, pem=-0.41162592,
similarity=26.385382]
36: 25391 Russia
    [pem_pos=1, final_scores=0.897345,
final_local_scores=0.94742936, pem=-1.0515022,
similarity=38.081127]
37: 4611 Burhanuddin_Rabbani
    [pem_pos=0, final_scores=0.3926034,
final_local_scores=0.4142449, pem=0.0, similarity=16.451607]
38: 25391 Russia
    [pem_pos=1, final_scores=0.6329023,
final_local_scores=0.65856594, pem=-1.8802736,
similarity=33.967854]
39: 69328 United_Arab_Emirates
    [pem_pos=0, final_scores=0.4592963,
final_local_scores=0.42787117, pem=-0.1950423,
similarity=17.918911]
40: 18950756 Abu_Dhabi
    [pem_pos=0, final_scores=0.2993161,
final_local_scores=0.28848568, pem=-0.065649725,
similarity=13.055715]
1: 18950756 Abu_Dhabi
    [pem_pos=0, final_scores=0.19111474,
final_local_scores=0.17287874, pem=-0.065649725,
similarity=9.605773]
```

Sample: b'17_Iraq'

[1 1 Iraq 1] 's [2 2 Saddam 2] meets [3 3 Russia 3] 's [4 4 Z
hirinovsky 4] . [5 5 BAGHDAD 5] 1996-08-22 [6 6 Iraqi 6]
President [7 7 Saddam Hussein 7] has told visiting [8 8 Russi
an 8] ultra-nationalist [9 9 Vladimir Zhirinovsky 9] that [10
10 Baghdad 10] wanted to maintain " friendship and
cooperation " with 1 Moscow , official [11 11 Iraqi 11]
newspapers said on Thursday . " President [12 12 Saddam Husse
in 12] stressed during the meeting [13 13 Iraq 13] 's
keenness to maintain friendship and cooperation with [14 14 Ru
ssia 14] , " the papers said . They said [15 15 Zhirinovsky 15
] told [16 16 Saddam 16] before he left [17 17 Baghdad 17] on
Wednesday that his Liberal Democratic party and the [18 18 Ru
ssian 18] [19 19 Duma 19] (parliament) " are calling for an
immediate lifting of the embargo " imposed on [20 20 Iraq 20]
after its 1990 invasion of [21 21 Kuwait 21] . [22 22 Zhirino
vsky 22] said on Tuesday he would press the [23 23 Russian 23]

government to help end [24 24 U.N. 24] trade sanctions on [25 25 Iraq 25] and blamed 2 Moscow for delaying establishment of good ties with [26 26 Baghdad 26] . " Our stand is firm , namely we are calling on (the [27 27 Russian 27]) government to end the economic embargo on [28 28 Iraq 28] and resume trade ties between [29 29 Russia 29] and [30 30 Iraq 30] , " he told reporters . [31 31 Zhirinovsky 31] visited [32 3 2 Iraq 32] twice in 1995 . Last October he was invited to attend the referendum held on [33 33 Iraq 33] 's presidency , which extended [34 34 Saddam 34] 's term for seven more years .

1: 7515928 Iraq
2: 29490 Saddam_Hussein
3: 25391 Russia
4: 392566 Vladimir_Zhirinovsky
5: 4492 Baghdad
6: 7515928 Iraq
7: 29490 Saddam_Hussein
8: 25391 Russia
9: 392566 Vladimir_Zhirinovsky
10: 7515928 Iraq
11: 7515928 Iraq
12: 29490 Saddam_Hussein
13: 7515928 Iraq
14: 25391 Russia
15: 392566 Vladimir_Zhirinovsky
16: 29490 Saddam_Hussein
17: 4492 Baghdad
18: 25391 Russia
19: 82288 Duma
20: 7515928 Iraq
21: 7515890 Kuwait
22: 392566 Vladimir_Zhirinovsky
23: 25391 Russia
24: 31769 United_Nations
25: 7515928 Iraq
26: 7515928 Iraq
27: 25391 Russia
28: 7515928 Iraq
29: 25391 Russia
30: 7515928 Iraq
31: 392566 Vladimir_Zhirinovsky
32: 7515928 Iraq
33: 7515928 Iraq
34: 29490 Saddam_Hussein
1: 25391 Russia
[winner: 19004 Moscow, pem_pos=0,
final_scores=-0.42925823, final_local_scores=-0.49952847,
pem=-0.106582455, similarity=-10.2376]
RECALL MISS: 25391 Russia]
2: 25391 Russia
[winner: 19004 Moscow, pem_pos=0,

final_scores=-0.33594358, final_local_scores=-0.3975964,
pem=-0.106582455, similarity=-7.1957445 |
RECALL MISS: 25391 Russia]

1: 7515928 Iraq
[pem_pos=0, final_scores=0.5910671,
final_local_scores=0.5418391, pem=-0.25379938,
similarity=21.639473]

2: 29490 Saddam_Hussein
[pem_pos=0, final_scores=0.43558592,
final_local_scores=0.41097802, pem=0.0, similarity=16.354116]

3: 25391 Russia
[pem_pos=0, final_scores=0.81757885,
final_local_scores=0.84769887, pem=-0.20299491,
similarity=30.490654]

4: 392566 Vladimir_Zhirinovsky
[pem_pos=0, final_scores=0.6073991,
final_local_scores=0.65420246, pem=0.0, similarity=23.61242]

5: 4492 Baghdad
[pem_pos=0, final_scores=0.4636129,
final_local_scores=0.43783748, pem=0.0, similarity=17.155657]

6: 7515928 Iraq
[pem_pos=0, final_scores=0.54651207,
final_local_scores=0.49316955, pem=-0.48469657,
similarity=21.442732]

7: 29490 Saddam_Hussein
[pem_pos=0, final_scores=0.48006248,
final_local_scores=0.45956188, pem=0.0, similarity=17.803957]

8: 25391 Russia
[pem_pos=1, final_scores=0.80532104,
final_local_scores=0.8343091, pem=-1.0515022,
similarity=34.70539]

9: 392566 Vladimir_Zhirinovsky
[pem_pos=0, final_scores=0.8213269,
final_local_scores=0.8878861, pem=0.0, similarity=30.586004]

10: 7515928 Iraq
[pem_pos=8, final_scores=0.2667523,
final_local_scores=0.1875745, pem=-2.5644467,
similarity=23.63317]

11: 7515928 Iraq
[pem_pos=0, final_scores=0.5433329,
final_local_scores=0.4896968, pem=-0.48469657,
similarity=21.339098]

12: 29490 Saddam_Hussein
[pem_pos=0, final_scores=0.52467084,
final_local_scores=0.50828975, pem=0.0, similarity=19.258093]

13: 7515928 Iraq
[pem_pos=0, final_scores=0.654411,
final_local_scores=0.6110328, pem=-0.25379938,
similarity=23.704351]

14: 25391 Russia
[pem_pos=0, final_scores=0.6764911,
final_local_scores=0.6935819, pem=-0.20299491,

```
similarity=25.891497]
15: 392566 Vladimir_Zhirinovsky
    [pem_pos=0, final_scores=0.6654805,
final_local_scores=0.71764755, pem=0.0, similarity=25.505747]
16: 29490 Saddam_Hussein
    [pem_pos=0, final_scores=0.42207986,
final_local_scores=0.39622468, pem=0.0, similarity=15.913848]
17: 4492 Baghdad
    [pem_pos=0, final_scores=0.37841046,
final_local_scores=0.34476674, pem=0.0, similarity=14.37824]
18: 25391 Russia
    [pem_pos=1, final_scores=0.83233184,
final_local_scores=0.8638143, pem=-1.0515022,
similarity=35.585888]
19: 82288 Duma
    [pem_pos=0, final_scores=0.32184786,
final_local_scores=0.3539229, pem=-0.19039723,
similarity=15.686889]
20: 7515928 Iraq
    [pem_pos=0, final_scores=0.60178226,
final_local_scores=0.5535438, pem=-0.25379938,
similarity=21.988766]
21: 7515890 Kuwait
    [pem_pos=0, final_scores=0.4603734,
final_local_scores=0.43173015, pem=-0.3372397,
similarity=18.807364]
22: 392566 Vladimir_Zhirinovsky
    [pem_pos=0, final_scores=0.6986087,
final_local_scores=0.7538351, pem=0.0, similarity=26.585655]
23: 25391 Russia
    [pem_pos=1, final_scores=0.8873628,
final_local_scores=0.92392725, pem=-1.0515022,
similarity=37.37978]
24: 31769 United_Nations
    [pem_pos=0, final_scores=0.48702592,
final_local_scores=0.50596416, pem=0.0, similarity=19.188692]
25: 7515928 Iraq
    [pem_pos=0, final_scores=0.63252836,
final_local_scores=0.5871293, pem=-0.25379938,
similarity=22.991022]
26: 7515928 Iraq
    [pem_pos=8, final_scores=0.26715446,
final_local_scores=0.18801379, pem=-2.5644467,
similarity=23.646278]
27: 25391 Russia
    [pem_pos=1, final_scores=0.71865845,
final_local_scores=0.73964334, pem=-1.0515022,
similarity=31.880379]
28: 7515928 Iraq
    [pem_pos=0, final_scores=0.62975425,
final_local_scores=0.584099, pem=-0.25379938,
similarity=22.900593]
```

```
29: 25391 Russia
    [pem_pos=0, final_scores=0.80309874,
final_local_scores=0.8318816, pem=-0.20299491,
similarity=30.018635]
30: 7515928 Iraq
    [pem_pos=0, final_scores=0.65316755,
final_local_scores=0.6096745, pem=-0.25379938,
similarity=23.663816]
31: 392566 Vladimir_Zhirinovsky
    [pem_pos=0, final_scores=0.6350193,
final_local_scores=0.6843733, pem=0.0, similarity=24.512777]
32: 7515928 Iraq
    [pem_pos=0, final_scores=0.63973147,
final_local_scores=0.5949976, pem=-0.25379938,
similarity=23.225828]
33: 7515928 Iraq
    [pem_pos=0, final_scores=0.56247556,
final_local_scores=0.51060724, pem=-0.25379938,
similarity=20.707453]
34: 29490 Saddam_Hussein
    [pem_pos=0, final_scores=0.36356193,
final_local_scores=0.33230272, pem=0.0, similarity=14.006289]
```

Sample: b'18_PRESS'

PRESS DIGEST - [1 1 Iraq 1] - Aug 22 . [2 2 BAGHDAD 2]
1996-08-22 These are some of the leading stories in the
official [3 3 Iraqi 3] press on Thursday . [4 4 Reuters 4]
has not verified these stories and does not vouch for their
accuracy . THAWRA - [5 5 Iraq 5] 's President [6 6 Saddam Hus
sein 6] meets with chairman of the [7 7 Russian 7] liberal
democratic party [8 8 Vladimir Zhirinovsky 8] . - Turkish
foreign minister says Turkey will take part in the [9 9 Bagh
ad 9] trade fair that will be held in November . [10 10 IRAQ 1
0] - A shipload of 12 tonnes of rice arrives in [11 11 Umm Qa
sr 11] port in the [12 12 Gulf 12] .

```
1: 7515928 Iraq
2: 4492 Baghdad
3: 7515928 Iraq
4: 18998750 Reuters
5: 7515928 Iraq
6: 29490 Saddam_Hussein
7: 25391 Russia
8: 392566 Vladimir_Zhirinovsky
9: 4492 Baghdad
10: 7515928 Iraq
11: 200154 Umm_Qasr
12: 24761 Persian_Gulf
1: 7515928 Iraq
    [pem_pos=0, final_scores=0.5989695,
final_local_scores=0.54177177, pem=-0.25379938,
similarity=21.637465]
2: 4492 Baghdad
```

```
[pem_pos=0, final_scores=0.44438374,
final_local_scores=0.39614373, pem=0.0, similarity=15.911432]
3: 7515928 Iraq
    [pem_pos=0, final_scores=0.53651386,
final_local_scores=0.4735484, pem=-0.48469657,
similarity=20.857197]
4: 18998750 Reuters
    [pem_pos=0, final_scores=0.33193594,
final_local_scores=0.35685635, pem=0.0, similarity=14.739018]
5: 7515928 Iraq
    [pem_pos=0, final_scores=0.62270933,
final_local_scores=0.5677039, pem=-0.25379938,
similarity=22.411331]
6: 29490 Saddam_Hussein
    [pem_pos=0, final_scores=0.5278905,
final_local_scores=0.51578325, pem=0.0, similarity=19.481714]
7: 25391 Russia
    [pem_pos=1, final_scores=0.74951243,
final_local_scores=0.8233155, pem=-1.0515022,
similarity=34.377323]
8: 392566 Vladimir_Zhirinovsky
    [pem_pos=0, final_scores=0.57003635,
final_local_scores=0.6674648, pem=0.0, similarity=24.008194]
9: 4492 Baghdad
    [pem_pos=0, final_scores=0.3390246,
final_local_scores=0.28105482, pem=0.0, similarity=12.47695]
10: 7515928 Iraq
    [pem_pos=0, final_scores=0.42551845,
final_local_scores=0.35230285, pem=-0.25379938,
similarity=15.983333]
11: 200154 Umm_Qasr
    [pem_pos=0, final_scores=0.32542866,
final_local_scores=0.3438854, pem=0.0, similarity=14.351939]
12: 24761 Persian_Gulf
    [pem_pos=4, final_scores=0.32055604,
final_local_scores=0.29637763, pem=-2.1612167,
similarity=24.687246]
```

Sample: b'19_PRESS'

PRESS DIGEST - [1 1 Lebanon 1] - Aug 22 . [2 2 BEIRUT 2]
1996-08-22 These are the leading stories in the [3 3 Beirut 3]
press on Thursday . [4 4 Reuters 4] has not verified these
stories and does not vouch for their accuracy . [5 5 AN-NAHAR
5] - Confrontation is escalating between [6 6 Hezbollah 6]
and the government . - Prime Minister Hariri : [7 7 Israeli 7]
threats do no serve peace . [8 8 AS-SAFIR 8] - Parliament
Speaker Berri : [9 9 Israel 9] is preparing for war against [1
0 10 Syria 10] and [11 11 Lebanon 11] . - Parliamentary
battle in [12 12 Beirut 12] .. The three main lists have been
prepared . AL-ANWAR - Continued criticism of law violation
incidents -- which occurred in the Mount Lebanon elections
last Sunday . [13 13 AD-DIYAR 13] - Financial negotiations

between [14 14 Lebanon 14] and [15 15 Pakistan 15] . - Hariri
to step into the election battle with an incomplete list .
NIDA'A AL-WATAN - [16 16 Maronite 16] Patriarch Sfeir
expressed sorrow over the violations in Sunday ' elections .

1: 17771 Lebanon
2: 37428 Beirut
3: 37428 Beirut
4: 18998750 Reuters
5: 2180853 An-Nahar
6: 13919 Hezbollah
7: 9282173 Israel
8: 4131343 As-Safir
9: 9282173 Israel
10: 7515849 Syria
11: 17771 Lebanon
12: 37428 Beirut
13: 4273826 Ad-Diyar
14: 17771 Lebanon
15: 23235 Pakistan
16: 1812373 Maronite_Church

1: 17771 Lebanon
[pem_pos=0, final_scores=0.5517863,
final_local_scores=0.49673134, pem=-0.4097839,
similarity=21.141636]
2: 37428 Beirut
[pem_pos=0, final_scores=0.50401896,
final_local_scores=0.4565072, pem=-0.046614014,
similarity=17.966293]
3: 37428 Beirut
[pem_pos=0, final_scores=0.4768167,
final_local_scores=0.42679286, pem=-0.046614014,
similarity=17.079557]
4: 18998750 Reuters
[pem_pos=0, final_scores=0.3331805,
final_local_scores=0.34367362, pem=0.0, similarity=14.345619]
5: 2180853 An-Nahar
[pem_pos=0, final_scores=0.6542239,
final_local_scores=0.71459335, pem=0.0, similarity=25.414604]
6: 13919 Hezbollah
[pem_pos=0, final_scores=0.35317987,
final_local_scores=0.33503625, pem=0.0, similarity=14.087863]
7: 9282173 Israel
[pem_pos=0, final_scores=0.53688484,
final_local_scores=0.51422364, pem=-0.34790626,
similarity=21.32714]
8: 4131343 As-Safir
[pem_pos=0, final_scores=0.51656514,
final_local_scores=0.5558549, pem=0.0, similarity=20.677532]
9: 9282173 Israel
[pem_pos=0, final_scores=0.6678057,
final_local_scores=0.6572347, pem=-0.21279564,
similarity=24.860125]

```
10: 7515849 Syria
    [pem_pos=0, final_scores=0.6598938,
final_local_scores=0.6424893, pem=-0.26162508,
similarity=24.685635]
11: 17771 Lebanon
    [pem_pos=0, final_scores=0.63114244,
final_local_scores=0.5834158, pem=-0.4097839,
similarity=23.728472]
12: 37428 Beirut
    [pem_pos=0, final_scores=0.5206542,
final_local_scores=0.47467864, pem=-0.046614014,
similarity=18.508564]
13: 4273826 Ad-Diyar
    [pem_pos=0, final_scores=0.4312529,
final_local_scores=0.57512814, pem=0.0, similarity=21.252684]
14: 17771 Lebanon
    [pem_pos=0, final_scores=0.6075974,
final_local_scores=0.5576965, pem=-0.4097839,
similarity=22.960958]
15: 23235 Pakistan
    [pem_pos=0, final_scores=0.5636749,
final_local_scores=0.5952842, pem=-0.21944217,
similarity=23.047543]
16: 1812373 Maronite_Church
    [pem_pos=0, final_scores=0.29672933,
final_local_scores=0.3087921, pem=0.0, similarity=13.304686]
```

Sample: b'20_CME'
[1 1 CME 1] live and feeder cattle calls range mixed . [2 2 C
HICAGO 2] 1996-08-22 Early calls on [3 3 CME 3] live and
feeder cattle futures ranged from 0.200 cent higher to 0.100
lower , livestock analysts said . The continued strong tone
to cash cattle and beef markets should prompt further
support . Outlook for a bullish cattle-on-feed report is also
expected to lend support and prompt some bull spreading ,
analysts said . However , trade will likely be light and
prices could drift on evening up ahead of the report . Cash
markets are also expected to be quiet after the record amount
of feedlot cattle traded this week , they said .

```
1: 351631 Chicago_Mercantile_Exchange
2: 6886 Chicago
3: 351631 Chicago_Mercantile_Exchange
1: 351631 Chicago_Mercantile_Exchange
    [pem_pos=1, final_scores=0.6741506,
final_local_scores=0.659844, pem=-1.04227, similarity=29.4488]
2: 6886 Chicago
    [pem_pos=0, final_scores=0.5043319,
final_local_scores=0.61567444, pem=-0.20658566,
similarity=23.586111]
3: 351631 Chicago_Mercantile_Exchange
    [pem_pos=1, final_scores=0.705046,
final_local_scores=0.69359255, pem=-1.04227,
```

similarity=30.455923]

Sample: b'21_Kindercare'

Kindercare says debt buy to hit Q1 results . [1 1 MONTGOMERY 1] , [2 2 Ala 2] . 1996-08-22 [3 3 KinderCare Learning Centers Inc 3] said on Thursday that a debt buyback would mean an extraordinary loss of \$ 1.2 million in its fiscal 1997 first quarter . The company said that during the quarter , which began June 1 , it bought \$ 30 million par value of its outstanding 10-3/8 percent senior notes due 2001 . The notes were bought for \$ 31.5 million . Philip Maslowe , chief financial officer of the preschool and child care company , said the buyback " offered an opportunity to reduce the company 's weighted average interest costs and improve future cash flows and earnings . "

1: 57690 Montgomery,_Alabama

2: 303 Alabama

3: 2563851 KinderCare_Learning_Centers

1: 57690 Montgomery,_Alabama

[pem_pos=0, final_scores=0.41574734,
final_local_scores=0.41771203, pem=-1.8303614,
similarity=26.508863]

2: 303 Alabama

[pem_pos=7, final_scores=0.7682628,
final_local_scores=0.79918134, pem=-2.937321,
similarity=43.912487]

3: 2563851 KinderCare_Learning_Centers

[pem_pos=0, final_scores=0.32629412,
final_local_scores=0.44269383, pem=0.0, similarity=17.30058]

Sample: b'22_RESEARCH'

RESEARCH ALERT - [1 1 Lehman 1] starts [2 2 SNET 2] . -- [3 3 Lehman 3] analyst Blake Bath started Southern New England Telecommunications Corp with an outperform rating , his office said . -- The analyst set a 12-month price target of \$ 45 and a fiscal 1996 year earnings estimate of \$ 3.09 per share , his office said . -- The analyst also set an earnings estimate for the 1997 year , but the figure was not immediately available . -- Southern New England closed at 38-1/2 Wednesday . -- E. Auchard , Wall Street bureau , 212-859-1736

1: 885075 Lehman_Brothers

2: 340894 Southern_New_England_Telecommunications

3: 885075 Lehman_Brothers

1: 885075 Lehman_Brothers

[pem_pos=0, final_scores=0.3367421,
final_local_scores=0.33042878, pem=-1.3331716,
similarity=21.200361]

2: 340894 Southern_New_England_Telecommunications

[pem_pos=2, final_scores=0.4179086,
final_local_scores=0.52658004, pem=-0.97135055,
similarity=25.086266]

```
3: 885075 Lehman_Brothers
    [pem_pos=0, final_scores=0.31572336,
final_local_scores=0.307469, pem=-1.3331716,
similarity=20.515196]

Sample: b'23_Gateway'
Gateway Data Sciences Q2 net rises . [1 1 PHOENIX 1]
1996-08-22 Summary of Consolidated Financial Data ( In
Thousands , except per share data ) Six Months Ended Quarter
Ended Jul 31 , Jul 31 , Jul 31 , Jul 31 , 1996 1995 1996 1995
Income Statement Data : Total Revenue $ 10,756 $ 13,102 $
7,961 $ 5,507 Software Revenue 2,383 1,558 1,086 1,074
Services Revenue 1,154 692 624 465 Operating Income 906 962
599 515 Net Income 821 512 565 301 Earnings Per Share 0.31
0.34 0.19 0.20 Jul 31 , 1996 Jan 31 , 1996 Balance Sheet
Data : Working Capital $ 5,755 ( $ 881 ) Cash and Cash
Equivalents 2,386 93 Total Assets 14,196 7,138 Shareholders '
Equity 5,951 ( 1,461 )
1: 49121 Phoenix,_Arizona
1: 49121 Phoenix,_Arizona
    [pem_pos=0, final_scores=0.18282278,
final_local_scores=0.47229332, pem=-0.90867877,
similarity=23.125425]

Sample: b'24_Greek'
[1 1 Greek 1] socialists give PM green light for election . [
2 2 ATHENS 2] 1996-08-22 The [3 3 Greek 3] socialist party 's
executive bureau gave Prime Minister [4 4 Costas Simitis 4]
its backing if he chooses to call snap elections , its
general secretary [5 5 Costas Skandalidis 5] told reporters
on Thursday . Prime Minister [6 6 Costas Simitis 6] will make
an official announcement after a cabinet meeting later on
Thursday , said Skandalidis . -- Dimitris Kontogiannis ,
Athens Newsroom +301 3311812-4
1: 12108 Greece
2: 1216 Athens
3: 12108 Greece
4: 393824 Costas_Simitis
5: 13871790 Kostas_Skandalidis
6: 393824 Costas_Simitis
1: 12108 Greece
    [pem_pos=1, final_scores=0.8402794,
final_local_scores=0.8592305, pem=-2.3247876,
similarity=42.37342]
2: 1216 Athens
    [pem_pos=0, final_scores=0.43985718,
final_local_scores=0.44993192, pem=-0.28831622,
similarity=19.084488]
3: 12108 Greece
    [pem_pos=1, final_scores=0.7495059,
final_local_scores=0.76007426, pem=-2.3247876,
similarity=39.414402]
```

```
4: 393824 Costas_Simitis
    [pem_pos=0, final_scores=0.6752767,
final_local_scores=0.7072681, pem=0.0, similarity=25.196005]
5: 13871790 Kostas_Skandalidis
    [pem_pos=0, final_scores=0.45506114,
final_local_scores=0.533776, pem=0.0, similarity=20.018654]
6: 393824 Costas_Simitis
    [pem_pos=0, final_scores=0.6377707,
final_local_scores=0.66629857, pem=0.0, similarity=23.97339]
```

Sample: b'25_PRESS'

PRESS DIGEST - [1 1 France 1] - [2 2 Le Monde 2] Aug 22 . [3 3 PARIS 3] 1996-08-22 These are leading stories in Thursday 's afternoon daily [4 4 Le Monde 4] , dated Aug 23 . FRONT PAGE -- [5 5 Africans 5] seeking to renew or obtain work and residence rights say Prime Minister [6 6 Alain Juppe 6] 's proposals are insufficient as hunger strike enters 49th day in [7 7 Paris 7] church and Wednesday rally attracts 8,000 sympathisers . -- [8 8 FLNC 8] [9 9 Corsican 9] nationalist movement announces end of truce after last night 's attacks . BUSINESS PAGES -- Shutdown of [10 10 Bally 10] 's [11 11 French 11] factories points up shoe industry crisis , with [12 12 French 12] manufacturers undercut by low-wage country competition and failure to keep abreast of trends . -- Secretary general of the [13 13 Sud-PTT 13] trade union at [14 14 France Telecom 14] all the elements are in place for social unrest in the next few weeks . -- Paris Newsroom +33 1 42 21 53 81

```
1: 5843419 France
2: 127071 Le_Monde
3: 22989 Paris
4: 127071 Le_Monde
5: 5334607 Africa
6: 434931 Alain_Juppé
7: 22989 Paris
8: 786920 National_Liberation_Front_of_Corsica
9: 5714828 Corsica
10: 2073198 Bally_Shoe
11: 5843419 France
12: 5843419 France
13: 3969921 Sud-PTT
14: 54314 Orange_S.A.
1: 5843419 France
    [pem_pos=0, final_scores=0.5516608,
final_local_scores=0.5160789, pem=-0.16240834,
similarity=20.373737]
2: 127071 Le_Monde
    [pem_pos=0, final_scores=0.47760004,
final_local_scores=0.5031125, pem=0.0, similarity=19.103594]
3: 22989 Paris
    [pem_pos=0, final_scores=0.5786794,
final_local_scores=0.55967945, pem=-0.115301244,
```

```
similarity=21.41869]
4: 127071 Le_Monde
    [pem_pos=0, final_scores=0.4629717,
final_local_scores=0.48713326, pem=0.0, similarity=18.626741]
5: 5334607 Africa
    [pem_pos=1, final_scores=0.58668655,
final_local_scores=0.65876096, pem=-1.1877377,
similarity=30.207558]
6: 434931 Alain_Juppé
    [pem_pos=0, final_scores=0.6891965,
final_local_scores=0.7716047, pem=0.0, similarity=27.115936]
7: 22989 Paris
    [pem_pos=0, final_scores=0.6561493,
final_local_scores=0.64430356, pem=-0.115301244,
similarity=23.944042]
8: 786920 National_Liberation_Front_of_Corsica
    [pem_pos=0, final_scores=0.48607308,
final_local_scores=0.5896018, pem=0.0, similarity=21.684607]
9: 5714828 Corsica
    [pem_pos=1, final_scores=0.39264423,
final_local_scores=0.41242987, pem=-1.0633177,
similarity=22.17993]
10: 2073198 Bally_Shoe
    [pem_pos=3, final_scores=0.30823392,
final_local_scores=0.43509942, pem=-1.8976667,
similarity=27.393753]
11: 5843419 France
    [pem_pos=0, final_scores=0.71759725,
final_local_scores=0.6973393, pem=-0.80223274,
similarity=29.262375]
12: 5843419 France
    [pem_pos=0, final_scores=0.77693826,
final_local_scores=0.76216024, pem=-0.80223274,
similarity=31.19676]
13: 3969921 Sud-PTT
    [pem_pos=0, final_scores=0.46516514,
final_local_scores=0.5928101, pem=0.0, similarity=21.780348]
14: 54314 Orange_S.A.
    [pem_pos=0, final_scores=0.27003723,
final_local_scores=0.3352002, pem=0.0, similarity=14.092755]
```

Sample: b'26_Well'

Well repairs to lift [1 1 Heidrun 1] oil output - [2 2 Statoil 2] . [3 3 OSL0 3] 1996-08-22 Three plugged water injection wells on the [4 4 Heidrun 4] oilfield off mid-Norway will be reopened over the next month , operator Den Norske Stats Oljeselskap AS ([5 5 Statoil 5]) said on Thursday . The plugged wells have accounted for a dip of 30,000 barrels per day (bpd) in [6 6 Heidrun 6] output to roughly 220,000 bpd , according to the company 's Status Weekly newsletter . The wells will be reperforated and gravel will be pumped into the reservoir through one of the wells to avoid plugging

```
problems in the future , it said . -- [7 7 Oslo 7] newsroom
+47 22 42 50 41
1: 2594976 Heidrun_oil_field
2: 11866976 Statoil
3: 22309 Oslo
4: 2594976 Heidrun_oil_field
5: 11866976 Statoil
6: 2594976 Heidrun_oil_field
7: 22309 Oslo
1: 2594976 Heidrun_oil_field
    [pem_pos=1, final_scores=0.6448229,
final_local_scores=0.685624, pem=-0.6095859,
similarity=27.865124]
2: 11866976 Statoil
    [pem_pos=0, final_scores=0.3555975,
final_local_scores=0.3583644, pem=0.0, similarity=14.784021]
3: 22309 Oslo
    [pem_pos=0, final_scores=0.49700373,
final_local_scores=0.5594742, pem=-0.035297178,
similarity=20.97749]
4: 2594976 Heidrun_oil_field
    [pem_pos=1, final_scores=0.4420237,
final_local_scores=0.46409667, pem=-0.6095859,
similarity=21.254307]
5: 11866976 Statoil
    [pem_pos=0, final_scores=0.3772905,
final_local_scores=0.3820607, pem=0.0, similarity=15.491166]
6: 2594976 Heidrun_oil_field
    [pem_pos=1, final_scores=0.5190882,
final_local_scores=0.548278, pem=-0.6095859,
similarity=23.766443]
7: 22309 Oslo
    [pem_pos=0, final_scores=0.5164715,
final_local_scores=0.5807398, pem=-0.035297178,
similarity=21.612097]
```

Sample: b'27_Finnish'

```
[1 1 Finnish 1] April trade surplus 3.8 billion markka -
NCB . [2 2 HELSINKI 2] 1996-08-22 [3 3 Finland 3] 's trade
surplus rose to 3.83 billion markka in April from 3.43
billion in March , the National Customs Board ( NCB ) said in
a statement on Thursday . The value of exports fell one
percent year-on-year in April and the value of imports fell
two percent , NCB said . Trade balance ( million markka ) :
April ' 96 March ' 96 Jan-April ' 96 Jan-April ' 95 Imports
10,663 10,725 43,430 40,989 Exports 14,494 14,153 56,126
56,261 Balance +3,831 +3,428 +12,696 +15,272 The January-
April 1995 import figure was revised from 39,584 million
markka and the export figure from 55,627 million markka . The
[4 4 Bank of Finland 4] earlier estimated the April trade
surplus at 3.2 billion markka with exports projected at 14.5
billion and imports at 11.3 billion . The NCB 's official
```

monthly trade statistics are lagging behind due to changes in customs procedures when [5 5 Finland 5] joined the [6 6 European Union 6] at the start of 1995 . -- Helsinki Newsroom +358 - 0 - 680 50 245

1: 10577 Finland

2: 13696 Helsinki

3: 10577 Finland

4: 850543 Bank_of_Finland

5: 10577 Finland

6: 9317 European_Union

1: 10577 Finland

[pem_pos=0, final_scores=0.681665, final_local_scores=0.6387239, pem=-0.8112885, similarity=27.56242]

2: 13696 Helsinki

[pem_pos=0, final_scores=0.6184468, final_local_scores=0.6080713, pem=-0.015504186, similarity=22.32009]

3: 10577 Finland

[pem_pos=0, final_scores=0.6682202, final_local_scores=0.6240376, pem=-0.2445894, similarity=24.042356]

4: 850543 Bank_of_Finland

[pem_pos=0, final_scores=0.59083027, final_local_scores=0.6910596, pem=0.0, similarity=24.712309]

5: 10577 Finland

[pem_pos=0, final_scores=0.633457, final_local_scores=0.586064, pem=-0.2445894, similarity=22.909147]

6: 9317 European_Union

[pem_pos=0, final_scores=0.39425665, final_local_scores=0.43493533, pem=-0.005814796, similarity=17.100674]

Sample: b'28_Dutch'

[1 1 Dutch 1] state raises tap sale price to 99.95 . [2 2 AM STERDAM 2] 1996-08-22 The Finance Ministry raised the price for tap sales of the [3 3 Dutch 3] government 's new 5.75 percent bond due September 2002 to 99.95 from 99.90 . Tap sales began on Monday and are being held daily from 07.00 [4 4 GMT 4] to 15.00 [5 5 GMT 5] until further notice . The ministry had raised 2.3 billion guilders from sales of the new bond by the close of trade on Wednesday . -- [6 6 Amsterd am 6] newsroom +31 20 504 5000

1: 21148 Netherlands

2: 844 Amsterdam

3: 21148 Netherlands

4: 12701 Greenwich_Mean_Time

5: 12701 Greenwich_Mean_Time

6: 844 Amsterdam

1: 21148 Netherlands

[pem_pos=0, final_scores=0.5631137,

```
final_local_scores=0.5324535, pem=-0.7943452,
similarity=24.29896]
2: 844 Amsterdam
    [pem_pos=0, final_scores=0.6369645,
final_local_scores=0.6227015, pem=-0.0642803,
similarity=23.021935]
3: 21148 Netherlands
    [pem_pos=0, final_scores=0.56892926,
final_local_scores=0.5388061, pem=-0.7943452,
similarity=24.488533]
4: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.6269031,
final_local_scores=0.6935865, pem=0.0, similarity=24.787718]
5: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.6202281,
final_local_scores=0.6862951, pem=0.0, similarity=24.570127]
6: 844 Amsterdam
    [pem_pos=0, final_scores=0.60288733,
final_local_scores=0.5854774, pem=-0.0642803,
similarity=21.911093]
```

Sample: b'29_German'

```
[1 1 German 1] farm ministry tells consumers to avoid [2 2 Br
itish 2] mutton . [3 3 BONN 3] 1996-08-22 [4 4 Germany 4] 's
Agriculture Ministry suggested on Wednesday that consumers
avoid eating meat from [5 5 British 5] sheep until scientists
determine whether mad cow disease can be transmitted to the
animals . " Until this is cleared up by the [6 6 European Uni
on 6] 's scientific panels -- and we have asked this to be
done as quickly as possible -- ( consumers ) should if at all
possible give preference to sheepmeat from other countries ,
" ministry official Werner Zwingmann told [7 7 ZDF 7]
television . " I do not want to say that there is a concrete
danger for consumers , " he added . " There are too many
holes in what we know , and these must be filled very
quickly . " Bonn has led efforts to ensure consumer
protection tops the list of priorities in dealing with the
mad cow crisis , which erupted in March when [8 8 Britain 8]
acknowledged humans could contract a similar illness by
eating contaminated beef . The [9 9 European Commission 9]
agreed this month to rethink a proposal to ban the use of
suspect sheep tissue after some EU veterinary experts
questioned whether it was justified . EU Farm Commissioner [10
10 Franz Fischler 10] had proposed banning sheep brains ,
spleens and spinal cords from the human and animal food
chains after reports from [11 11 Britain 11] and [12 12 France
12] that under laboratory conditions sheep could contract
Bovine Spongiform Encephalopathy ( [13 13 BSE 13] ) -- mad
cow disease . But some members of the EU 's standing
veterinary committee questioned whether the action was
necessary given the slight risk to human health . The
question is being studied separately by two EU scientific
```

committees . Sheep have long been known to contract scrapie , a similar brain-wasting disease to [14 14 BSE 14] which is believed to have been transferred to cattle through feed containing animal waste . [15 15 British 15] officials say sheep meat is perfectly safe to eat . [16 16 ZDF 16] said [17 17 Germany 17] imported 47,600 sheep from [18 18 Britain 18] last year , nearly half of total imports . It brought in 4,275 tonnes of [19 19 British 19] mutton , some 10 percent of overall imports . After the [20 20 British 20] government admitted a possible link between mad cow disease and its fatal human equivalent , the EU imposed a worldwide ban on [21 21 British 21] beef exports . EU leaders agreed at a summit in June to a progressive lifting of the ban as [22 22 Britain 22] takes parallel measures to eradicate the disease .

1: 11867 Germany
2: 31717 United_Kingdom
3: 3295 Bonn
4: 11867 Germany
5: 31717 United_Kingdom
6: 9317 European_Union
7: 262201 ZDF
8: 31717 United_Kingdom
9: 9974 European_Commission
10: 626779 Franz_Fischler
11: 31717 United_Kingdom
12: 5843419 France
13: 19344418 Bovine_spongiform_encephalopathy
14: 19344418 Bovine_spongiform_encephalopathy
15: 31717 United_Kingdom
16: 262201 ZDF
17: 11867 Germany
18: 31717 United_Kingdom
19: 31717 United_Kingdom
20: 31717 United_Kingdom
21: 31717 United_Kingdom
22: 31717 United_Kingdom
1: 11867 Germany
 [pem_pos=1, final_scores=0.80958444,
final_local_scores=0.8096456, pem=-1.0465823,
similarity=33.94263]
2: 31717 United_Kingdom
 [pem_pos=0, final_scores=0.6276752,
final_local_scores=0.57748634, pem=-0.48936236,
similarity=23.984285]
3: 3295 Bonn
 [pem_pos=0, final_scores=0.4689259,
final_local_scores=0.47414285, pem=0.0, similarity=18.23908]
4: 11867 Germany
 [pem_pos=0, final_scores=0.8561632,
final_local_scores=0.8605258, pem=-0.15529114,
similarity=30.614016]
5: 31717 United_Kingdom

```
[pem_pos=0, final_scores=0.6167398,
final_local_scores=0.56554115, pem=-0.48936236,
similarity=23.627817]
6: 9317 European_Union
    [pem_pos=0, final_scores=0.45804304,
final_local_scores=0.45614898, pem=-0.005814796,
similarity=17.73373]
7: 262201 ZDF
    [pem_pos=0, final_scores=0.38691133,
final_local_scores=0.4190626, pem=0.0, similarity=16.595377]
8: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5540152,
final_local_scores=0.497024, pem=-0.61114657,
similarity=22.24541]
9: 9974 European_Commission
    [pem_pos=0, final_scores=0.44476634,
final_local_scores=0.46717703, pem=0.0, similarity=18.031208]
10: 626779 Franz_Fischler
    [pem_pos=0, final_scores=0.56328213,
final_local_scores=0.7289707, pem=0.0, similarity=25.843653]
11: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5787026,
final_local_scores=0.5239913, pem=-0.61114657,
similarity=23.050167]
12: 5843419 France
    [pem_pos=0, final_scores=0.691732,
final_local_scores=0.695006, pem=-0.16240834,
similarity=25.71328]
13: 19344418 Bovine_spongiform_encephalopathy
    [pem_pos=0, final_scores=0.58832914,
final_local_scores=0.6949407, pem=-0.2201173,
similarity=26.02516]
14: 19344418 Bovine_spongiform_encephalopathy
    [pem_pos=0, final_scores=0.689753,
final_local_scores=0.8057308, pem=-0.2201173,
similarity=29.33136]
15: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5230904,
final_local_scores=0.46324342, pem=-0.48936236,
similarity=20.57505]
16: 262201 ZDF
    [pem_pos=0, final_scores=0.39690465,
final_local_scores=0.4299788, pem=0.0, similarity=16.921137]
17: 11867 Germany
    [pem_pos=0, final_scores=0.68460983,
final_local_scores=0.6731298, pem=-0.15529114,
similarity=25.021746]
18: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.58531183,
final_local_scores=0.53121084, pem=-0.61114657,
similarity=23.265614]
19: 31717 United_Kingdom
```

```
[pem_pos=0, final_scores=0.56077045,
final_local_scores=0.50440305, pem=-0.48936236,
similarity=21.803335]
20: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.6117629,
final_local_scores=0.56010455, pem=-0.48936236,
similarity=23.465578]
21: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.50518066,
final_local_scores=0.4436797, pem=-0.48936236,
similarity=19.991228]
22: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5490305,
final_local_scores=0.491579, pem=-0.61114657,
similarity=22.082922]
```

Sample: b'30_GOLF'

GOLF - SCORES AT [1 1 WORLD SERIES OF GOLF 1] . [2 2 AKRON 2] , [3 3 Ohio 3] 1996-08-22 Scores from the \$ 2.1 million [4 4 N EC World Series of Golf 4] after the first round Thursday at the 7,149 yard , par 70 Firestone C.C course (players [5 5 U. S. 5] unless stated) : 66 [6 6 Paul Goydos 6] , [7 7 Billy Mayfair 7] , [8 8 Hidemichi Tanaka 8] ([9 9 Japan 9]) 68 [10 1 0 Steve Stricker 10] 69 [11 11 Justin Leonard 11] , [12 12 Mark Brooks 12] 70 [13 13 Tim Herron 13] , [14 14 Duffy Waldorf 14] , [15 15 Davis Love 15] , [16 16 Anders Forsbrand 16] ([1 7 17 Sweden 17]) , [18 18 Nick Faldo 18] ([19 19 Britain 19]) , [20 20 John Cook 20] , [21 21 Steve Jones 21] , [22 22 Phil Mickelson 22] , [23 23 Greg Norman 23] ([24 24 Australia 24]) 71 [25 25 Ernie Els 25] ([26 26 South Africa 26]) , [2 7 27 Scott Hoch 27] 72 [28 28 Clarence Rose 28] , [29 29 Loren Roberts 29] , [30 30 Fred Funk 30] , [31 31 Sven Struver 31] ([32 32 Germany 32]) , [33 33 Alexander Cejka 33] ([34 34 Germany 34]) , [35 35 Hal Sutton 35] , [36 36 Tom Lehman 36] 73 [37 37 D.A. Weibring 37] , [38 38 Brad Bryant 38] , [39 39 Craig Parry 39] ([40 40 Australia 40]) , [41 41 Stewart Cink 41] ([42 42 Australia 42]) , [43 43 Corey Pavin 43] , [44 44 Craig Stadler 44] , Mark O'Meara , [45 45 Fred Couples 45] 74 [46 46 Paul Stankowski 46] , [47 47 Costantino Rocca 47] ([48 48 Italy 48]) 75 [49 49 Jim Furyk 49] , [50 50 Satoshi Higashii 50] ([51 51 Japan 51]) , [52 52 Willie Wood 52] , [53 53 Shigeki Maruyama 53] ([54 54 Japan 54]) 76 [55 55 Scott McCarron 55] 77 [56 56 Wayne Westner 56] ([57 57 South Africa 57]) , Steve Schneiter 79 [58 58 Tom Watson 58] 81 Seiki Okuda ([59 59 Japan 59])

```
1: 1536678 World_Series_of_Golf
2: 129842 Akron,_Ohio
3: 22199 Ohio
4: 1527695 WGC-Bridgestone_Invitational
5: 3434750 United_States
6: 6280981 Paul_Goydos
7: 2080225 Billy_Mayfair
```

8: 2114665 Hidemichi_Tanaka
9: 15573 Japan
10: 2310551 Steve_Stricker
11: 1507170 Justin_Leonard
12: 2295891 Mark_Brooks_(golfer)
13: 5242796 Tim_Herron
14: 1781343 Duffy_Waldorf
15: 1388818 Davis_Love_III
16: 4377636 Anders_Forsbrand
17: 5058739 Sweden
18: 759285 Nick_Faldo
19: 31717 United_Kingdom
20: 2810777 John_Cook_(golfer)
21: 1924538 Steve_Jones_(golfer)
22: 592392 Phil_Mickelson
23: 157447 Greg_Norman
24: 4689264 Australia
25: 230467 Ernie_Els
26: 17416221 South_Africa
27: 3396339 Scott_Hoch
28: 8385080 Clarence_Rose
29: 2567564 Loren_Roberts
30: 1646050 Fred_Funk
31: 4337089 Sven_Strüver
32: 11867 Germany
33: 2080183 Alex_Čejka
34: 11867 Germany
35: 654136 Hal_Sutton
36: 1516564 Tom_Lehman
37: 5653062 D._A._Weibring
38: 4450518 Brad_Bryant
39: 3738086 Craig_Parry
40: 4689264 Australia
41: 5418317 Stewart_Ginn
42: 4689264 Australia
43: 1530024 Corey_Pavin
44: 1261748 Craig_Stadler
45: 1221065 Fred_Couples
46: 6189303 Paul_Stankowski
47: 2157151 Costantino_Rocca
48: 14532 Italy
49: 247955 Jim_Furyk
50: 21138421 Satoshi_Higashi
51: 15573 Japan
52: 7018258 Willie_Wood_(golfer)
53: 414220 Shigeki_Maruyama
54: 15573 Japan
55: 5845032 Scott_McCarron
56: 5206729 Wayne_Westner
57: 17416221 South_Africa
58: 585794 Tom_Watson_(golfer)
59: 15573 Japan

```
1: 1536678 World_Series_of_Golf
    [pem_pos=0, final_scores=0.42440033,
final_local_scores=0.49276555, pem=0.0, similarity=18.794819]
2: 129842 Akron,_Ohio
    [pem_pos=0, final_scores=0.55852616,
final_local_scores=0.6264395, pem=-0.6020413,
similarity=26.057913]
3: 22199 Ohio
    [pem_pos=0, final_scores=0.5921374,
final_local_scores=0.68345934, pem=-0.19147073,
similarity=25.52675]
4: 1527695 WGC-Bridgestone_Invitational
    [pem_pos=0, final_scores=0.40730822,
final_local_scores=0.42434675, pem=0.0, similarity=16.753067]
5: 3434750 United_States
    [pem_pos=0, final_scores=0.42814404,
final_local_scores=0.48041517, pem=-0.14075686,
similarity=19.19172]
6: 6280981 Paul_Goydos
    [pem_pos=0, final_scores=0.74082714,
final_local_scores=0.7879404, pem=0.0, similarity=27.603426]
7: 2080225 Billy_Mayfair
    [pem_pos=0, final_scores=0.71168363,
final_local_scores=0.7558206, pem=0.0, similarity=26.644907]
8: 2114665 Hidemichi_Tanaka
    [pem_pos=0, final_scores=0.47786254,
final_local_scores=0.61266077, pem=0.0, similarity=22.372732]
9: 15573 Japan
    [pem_pos=0, final_scores=0.6569054,
final_local_scores=0.74065197, pem=-0.13549899,
similarity=26.92911]
10: 2310551 Steve_Stricker
    [pem_pos=0, final_scores=0.71302253,
final_local_scores=0.7442846, pem=0.0, similarity=26.30065]
11: 1507170 Justin_Leonard
    [pem_pos=0, final_scores=0.78961915,
final_local_scores=0.8294123, pem=0.0, similarity=28.84103]
12: 2295891 Mark_Brooks_(golfer)
    [pem_pos=0, final_scores=0.63394,
final_local_scores=0.6768205, pem=-0.17831625,
similarity=25.257097]
13: 5242796 Tim_Herron
    [pem_pos=0, final_scores=0.605537,
final_local_scores=0.64851946, pem=0.0, similarity=23.442827]
14: 1781343 Duffy_Waldorf
    [pem_pos=0, final_scores=0.614575,
final_local_scores=0.6903748, pem=0.0, similarity=24.691874]
15: 1388818 Davis_Love_III
    [pem_pos=0, final_scores=0.62201905,
final_local_scores=0.6324935, pem=0.0, similarity=22.96458]
16: 4377636 Anders_Forsbrand
    [pem_pos=0, final_scores=0.5915697,
```

```
final_local_scores=0.65138483, pem=0.0, similarity=23.528336]
17: 5058739 Sweden
    [pem_pos=0, final_scores=0.68432426,
final_local_scores=0.76650447, pem=-0.2135508,
similarity=28.125057]
18: 759285 Nick_Faldo
    [pem_pos=0, final_scores=0.9395954,
final_local_scores=0.97745806, pem=0.0, similarity=33.25901]
19: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.50505596,
final_local_scores=0.5872875, pem=-0.61114657,
similarity=24.939053]
20: 2810777 John_Cook_(golfer)
    [pem_pos=1, final_scores=0.4998135,
final_local_scores=0.50455725, pem=-1.1074443,
similarity=25.16916]
21: 1924538 Steve_Jones_(golfer)
    [pem_pos=4, final_scores=0.41432327,
final_local_scores=0.43461496, pem=-2.187635,
similarity=28.956188]
22: 592392 Phil_Mickelson
    [pem_pos=0, final_scores=0.8420232,
final_local_scores=0.8723443, pem=0.0, similarity=30.122208]
23: 157447 Greg_Norman
    [pem_pos=0, final_scores=0.645434,
final_local_scores=0.65818685, pem=0.0, similarity=23.731321]
24: 4689264 Australia
    [pem_pos=0, final_scores=0.6523764,
final_local_scores=0.73067385, pem=-0.17168477,
similarity=26.828127]
25: 230467 Ernie_Els
    [pem_pos=0, final_scores=0.70958984,
final_local_scores=0.73769873, pem=0.0, similarity=26.104115]
26: 17416221 South_Africa
    [pem_pos=0, final_scores=0.55039626,
final_local_scores=0.63510865, pem=-0.26251382,
similarity=24.470213]
27: 3396339 Scott_Hoch
    [pem_pos=0, final_scores=0.72821146,
final_local_scores=0.7796382, pem=0.0, similarity=27.35567]
28: 8385080 Clarence_Rose
    [pem_pos=0, final_scores=0.51194304,
final_local_scores=0.59844285, pem=0.0, similarity=21.94844]
29: 2567564 Loren_Roberts
    [pem_pos=0, final_scores=0.60376984,
final_local_scores=0.62516737, pem=0.0, similarity=22.745955]
30: 1646050 Fred_Funk
    [pem_pos=0, final_scores=0.5344905,
final_local_scores=0.5558665, pem=0.0, similarity=20.677877]
31: 4337089 Sven_Strüver
    [pem_pos=0, final_scores=0.6986714,
final_local_scores=0.7983995, pem=0.0, similarity=27.915546]
```

32: 11867 Germany
[pem_pos=0, final_scores=0.69099355,
final_local_scores=0.7862003, pem=-0.15529114,
similarity=28.395996]
33: 2080183 Alex Čejka
[pem_pos=0, final_scores=0.8556381,
final_local_scores=0.9194172, pem=0.0, similarity=31.526957]
34: 11867 Germany
[pem_pos=0, final_scores=0.7280959,
final_local_scores=0.82672894, pem=-0.15529114,
similarity=29.605452]
35: 654136 Hal Sutton
[pem_pos=0, final_scores=0.5981392,
final_local_scores=0.6280239, pem=0.0, similarity=22.8312]
36: 1516564 Tom Lehman
[pem_pos=0, final_scores=0.90356725,
final_local_scores=0.93761116, pem=0.0, similarity=32.0699]
37: 5653062 D. A. Weibring
[pem_pos=0, final_scores=0.64873666,
final_local_scores=0.7069335, pem=0.0, similarity=25.186018]
38: 4450518 Brad Bryant
[pem_pos=0, final_scores=0.6370236,
final_local_scores=0.6811149, pem=0.0, similarity=24.41554]
39: 3738086 Craig Parry
[pem_pos=0, final_scores=0.6559529,
final_local_scores=0.68349224, pem=0.0, similarity=24.486485]
40: 4689264 Australia
[pem_pos=0, final_scores=0.7426357,
final_local_scores=0.82926846, pem=-0.17168477,
similarity=29.770388]
41: 5418317 Stewart Ginn
[pem_pos=0, final_scores=0.47812665,
final_local_scores=0.5787747, pem=0.0, similarity=21.361504]
42: 4689264 Australia
[pem_pos=0, final_scores=0.6787206,
final_local_scores=0.75945085, pem=-0.17168477,
similarity=27.68689]
43: 1530024 Corey Pavin
[pem_pos=0, final_scores=0.8761657,
final_local_scores=0.90876323, pem=0.0, similarity=31.20902]
44: 1261748 Craig Stadler
[pem_pos=0, final_scores=0.66651237,
final_local_scores=0.69679344, pem=0.0, similarity=24.883419]
45: 1221065 Fred Couples
[pem_pos=0, final_scores=0.72737277,
final_local_scores=0.7541875, pem=0.0, similarity=26.596172]
46: 6189303 Paul Stankowski
[pem_pos=0, final_scores=0.68280125,
final_local_scores=0.7682221, pem=0.0, similarity=27.014992]
47: 2157151 Costantino Rocca
[pem_pos=0, final_scores=0.65310425,
final_local_scores=0.69676596, pem=0.0, similarity=24.882599]

```
48: 14532 Italy
    [pem_pos=0, final_scores=0.6864596,
final_local_scores=0.8028796, pem=-0.20612453,
similarity=29.17018]
49: 247955 Jim_Furyk
    [pem_pos=0, final_scores=0.8194801,
final_local_scores=0.8489753, pem=0.0, similarity=29.42483]
50: 21138421 Satoshi_Higashi
    [pem_pos=0, final_scores=0.6288515,
final_local_scores=0.7480796, pem=0.0, similarity=26.4139]
51: 15573 Japan
    [pem_pos=0, final_scores=0.73761886,
final_local_scores=0.82881916, pem=-0.13549899,
similarity=29.560194]
52: 7018258 Willie_Wood_(golfer)
    [pem_pos=1, final_scores=0.5519143,
final_local_scores=0.5950963, pem=-0.86820155,
similarity=26.569988]
53: 414220 Shigeki_Maruyama
    [pem_pos=0, final_scores=0.6033116,
final_local_scores=0.65022415, pem=0.0, similarity=23.493698]
54: 15573 Japan
    [pem_pos=0, final_scores=0.717096,
final_local_scores=0.806401, pem=-0.13549899,
similarity=28.891191]
55: 5845032 Scott_McCarron
    [pem_pos=0, final_scores=0.6890774,
final_local_scores=0.75551903, pem=0.0, similarity=26.635908]
56: 5206729 Wayne_Westner
    [pem_pos=0, final_scores=0.74396807,
final_local_scores=0.83215773, pem=0.0, similarity=28.92296]
57: 17416221 South_Africa
    [pem_pos=0, final_scores=0.58952343,
final_local_scores=0.6778492, pem=-0.26251382,
similarity=25.745676]
58: 585794 Tom_Watson_(golfer)
    [pem_pos=0, final_scores=0.61157656,
final_local_scores=0.6252155, pem=-0.48845047,
similarity=25.40366]
59: 15573 Japan
    [pem_pos=0, final_scores=0.70387983,
final_local_scores=0.79196435, pem=-0.13549899,
similarity=28.460373]
```

Sample: b'31_SOCCER'

SOCCER - [1 1 GLORIA BISTRITA 1] BEAT 2-1 F.C. VALLETTA . [2 2 BISTRITA 2] 1996-08-22 [3 3 Gloria Bistrita 3] ([4 4 Roman ia 4]) beat 2-1 (halftime 1-1) F.C. Valletta ([5 5 Malta 5]) in their [6 6 Cup winners Cup 6] match , second leg of the preliminary round , on Thursday . Scorers : [7 7 Gloria Bi strita 7] - Ilie Lazar (32nd) , Eugen Voica (84th) F.C. La Valletta - [8 8 Gilbert Agius 8] (24th) Attendance :

8,000 [9 9 Gloria_Bistrita 9] won 4-2 on aggregate and qualified for the first round of the [10 10 Cup winners Cup 10] . REUTER

1: 2384775 ACF_Gloria_Bistrița
2: 415846 Bistrița
3: 2384775 ACF_Gloria_Bistrița
4: 25445 Romania
5: 19137 Malta
6: 88178 UEFA_Cup_Winners'_Cup
7: 2384775 ACF_Gloria_Bistrița
8: 10424084 Gilbert_Agius
9: 2384775 ACF_Gloria_Bistrița
10: 88178 UEFA_Cup_Winners'_Cup

1: 2384775 ACF_Gloria_Bistrița
[pem_pos=0, final_scores=0.58424306,
final_local_scores=0.5528128, pem=0.0, similarity=20.58675]
2: 415846 Bistrița
[pem_pos=0, final_scores=0.2680592,
final_local_scores=0.36811203, pem=0.0, similarity=15.074909]
3: 2384775 ACF_Gloria_Bistrița
[pem_pos=0, final_scores=0.70512736,
final_local_scores=0.6848605, pem=0.0, similarity=24.527317]
4: 25445 Romania
[pem_pos=0, final_scores=0.5561496,
final_local_scores=0.6246278, pem=-0.21241479,
similarity=23.884998]
5: 19137 Malta
[pem_pos=0, final_scores=0.22512273,
final_local_scores=0.3004638, pem=-0.39222723,
similarity=15.189144]
6: 88178 UEFA_Cup_Winners'_Cup
[pem_pos=0, final_scores=0.48313218,
final_local_scores=0.48581934, pem=-0.028414067,
similarity=18.74205]
7: 2384775 ACF_Gloria_Bistrița
[pem_pos=0, final_scores=0.94714075,
final_local_scores=0.94922346, pem=0.0, similarity=32.416435]
8: 10424084 Gilbert_Agius
[pem_pos=0, final_scores=0.5926989,
final_local_scores=0.7089908, pem=0.0, similarity=25.247414]
9: 2384775 ACF_Gloria_Bistrița
[pem_pos=0, final_scores=0.7590088,
final_local_scores=0.74371785, pem=0.0, similarity=26.283737]
10: 88178 UEFA_Cup_Winners'_Cup
[pem_pos=0, final_scores=0.4299898,
final_local_scores=0.42776936, pem=-0.028414067,
similarity=17.009724]

Sample: b'32_HORSE'

HORSE RACING - PIVOTAL ENDS 25-YEAR WAIT FOR TRAINER
PRESCOTT . [1 1 YORK 1] , [2 2 England 2] 1996-08-22 Sir [3 3
Mark Prescott 3] landed his first group one victory in 25

years as a trainer when his top sprinter Pivotal , a 100-30 chance , won the [4 4 Nunthorpe Stakes 4] on Thursday . The three-year-old , partnered by veteran [5 5 George Duffield 5] , snatched a short head verdict in the last stride to deny Eveningperformance (16-1) , trained by Henry Candy and ridden by Chris Rutter . Hever Golf Rose (11-4) , last year 's Prix de l ' Abbaye winner at [6 6 Longchamp 6] , finished third , a further one and a quarter lengths away with the 7-4 favourite Mind Games in fourth . Pivotal , a [7 7 Royal Ascot 7] winner in June , may now be aimed at this season 's Abbaye , [8 8 Europe 8] 's top sprint race . Prescott , reluctant to go into the winner 's enclosure until the result of the photo-finish was announced , said : " Twenty-five years and I have never been there so I thought I had better wait a bit longer . " He added : " It 's very sad to beat Henry Candy because I am godfather to his daughter . " Like Prescott , Jack Berry , trainer of Mind Games , had gone into Thursday 's race in search of a first group one success after many years around the top of his profession . Berry said : " I `m disappointed but I do n't feel suicidal . He (Mind Games) was going as well as any of them one and a half furlongs (300 metres) out but he just did n't quicken . "

1: 34361 York

2: 9316 England

3: 25674876 Mark_Prescott

4: 727606 Nunthorpe_Stakes

5: 7183198 George_Duffield

6: 1204250 Longchamp_Racecourse

7: 826248 Ascot_Racecourse

8: 9239 Europe

1: 34361 York

[pem_pos=0, final_scores=0.4881429,
final_local_scores=0.51616806, pem=-0.7238267,
similarity=23.42948]

2: 9316 England

[pem_pos=0, final_scores=0.6545985,
final_local_scores=0.7041408, pem=-0.19053856,
similarity=26.138857]

3: 25674876 Mark_Prescott

[pem_pos=0, final_scores=0.45870143,
final_local_scores=0.5090502, pem=0.0, similarity=19.280787]

4: 727606 Nunthorpe_Stakes

[pem_pos=0, final_scores=0.51706177,
final_local_scores=0.5670829, pem=0.0, similarity=21.012598]

5: 7183198 George_Duffield

[pem_pos=0, final_scores=0.48958284,
final_local_scores=0.57079923, pem=0.0, similarity=21.1235]

6: 1204250 Longchamp_Racecourse

[pem_pos=0, final_scores=0.38675618,
final_local_scores=0.42137212, pem=-0.26997083,

similarity=18.132442]

7: 826248 Ascot_Racecourse

```
[pem_pos=0, final_scores=0.5009513,
final_local_scores=0.49512506, pem=0.0, similarity=18.865232]
8: 9239 Europe
[pem_pos=0, final_scores=0.52417105,
final_local_scores=0.6521304, pem=-0.086338006,
similarity=24.020105]
```

Sample: b'33_HORSE'

HORSE RACING - NUNTHORPE STAKES RESULTS . [1 1 YORK 1] , [2 2 England 2] 1996-08-22 Result of the [3 3 Nunthorpe Stakes 3] , a group one race for two-year-olds and upwards , run over five furlongs (1 km) on Thursday : 1. Pivotal 100-30 (ridden by [4 4 George Duffield 4]) 2. Eveningperformance 16-1 (Chris Rutter) 3. Hever Golf Rose 11-4 ([5 5 Jason Weaver 5]) Eight ran . Favourite : Mind Games (7-4) finished 4th Distances : a short head , 1-1/4 lengths . Winner owned by the Cheveley Park Stud and trained by Sir [6 6 Mark Prescott 6] at [7 7 Newmarket 7] . Value to winner : 72,464 pounds sterling (\$ 112,200)

1: 34361 York

2: 9316 England

3: 727606 Nunthorpe_Stakes

4: 7183198 George_Duffield

5: 18983954 Jason_Weaver_(jockey)

6: 25674876 Mark_Prescott

7: 158390 Newmarket,_Suffolk

1: 34361 York

```
[pem_pos=0, final_scores=0.4838583,
final_local_scores=0.5204904, pem=-0.7238267,
similarity=23.558468]
```

2: 9316 England

```
[pem_pos=0, final_scores=0.6162314,
final_local_scores=0.6782396, pem=-0.19053856,
similarity=25.365913]
```

3: 727606 Nunthorpe_Stakes

```
[pem_pos=0, final_scores=0.51634675,
final_local_scores=0.55550313, pem=0.0, similarity=20.667034]
```

4: 7183198 George_Duffield

```
[pem_pos=0, final_scores=0.5414767,
final_local_scores=0.6101684, pem=0.0, similarity=22.298355]
```

5: 18983954 Jason_Weaver_(jockey)

```
[pem_pos=1, final_scores=0.3387612,
final_local_scores=0.4020331, pem=-0.21632794,
similarity=17.263607]
```

6: 25674876 Mark_Prescott

```
[pem_pos=0, final_scores=0.47975212,
final_local_scores=0.55871135, pem=0.0, similarity=20.762774]
```

7: 158390 Newmarket,_Suffolk

```
[pem_pos=0, final_scores=0.35344464,
final_local_scores=0.32331637, pem=-1.2625822,
similarity=20.604237]
```

```
Sample: b'34_TENNIS'
TENNIS - RESULTS AT [1 1 TOSHIBA CLASSIC 1] . [2 2 CARLSBAD 2
] , [3 3 California 3] 1996-08-21 Results from the $ 450,000 [
4 4 Toshiba Classic 4] tennis tournament on Wednesday
( prefix number denotes seeding ) : Second round 1 - [5 5 Ara
ntxa Sanchez Vicario 5] ( [6 6 Spain 6] ) beat Naoko Kijimuta
( [7 7 Japan 7] ) 1-6 6-4 6-3 4 - [8 8 Kimiko Date 8] ( [9 9 J
apan 9] ) beat Yone Kamio ( [10 10 Japan 10] ) 6-2 7-5 [11 11
Sandrine Testud 11] ( [12 12 France 12] ) beat 7 - [13 13 Ai S
ugiyama 13] ( [14 14 Japan 14] ) 6-3 4-6 6-4 8 - [15 15 Natha
lie Tauziat 15] ( [16 16 France 16] ) beat [17 17 Shi-Ting Wa
ng 17] ( [18 18 Taiwan 18] ) 6-4 6-2
1: 31726664 Southern_California_Open
2: 108030 Carlsbad,_California
3: 5407 California
4: 31726664 Southern_California_Open
5: 534624 Arantxa_Sánchez_Vicario
6: 26667 Spain
7: 15573 Japan
8: 495296 Kimiko_Date-Krumm
9: 15573 Japan
10: 15573 Japan
11: 3127711 Sandrine_Testud
12: 5843419 France
13: 1514559 Ai_Sugiyama
14: 15573 Japan
15: 1028366 Nathalie_Tauziat
16: 5843419 France
17: 15340120 Wang_Shing
18: 25734 Taiwan
1: 31726664 Southern_California_Open
    [pem_pos=0, final_scores=0.44486302,
final_local_scores=0.4936307, pem=0.0, similarity=18.820639]
2: 108030 Carlsbad,_California
    [pem_pos=0, final_scores=0.5019271,
final_local_scores=0.5638518, pem=-0.36353475,
similarity=22.893133]
3: 5407 California
    [pem_pos=0, final_scores=0.6206412,
final_local_scores=0.69035566, pem=-0.08662026,
similarity=25.162357]
4: 31726664 Southern_California_Open
    [pem_pos=0, final_scores=0.5154042,
final_local_scores=0.5706862, pem=0.0, similarity=21.120129]
5: 534624 Arantxa_Sánchez_Vicario
    [pem_pos=0, final_scores=0.55808634,
final_local_scores=0.62555224, pem=0.0, similarity=22.757439]
6: 26667 Spain
    [pem_pos=0, final_scores=0.603588,
final_local_scores=0.6363146, pem=-0.18857051,
similarity=24.104084]
7: 15573 Japan
```

```
[pem_pos=0, final_scores=0.78871924,
final_local_scores=0.82149476, pem=-0.13549899,
similarity=29.34162]
8: 495296 Kimiko_Date-Krumm
    [pem_pos=0, final_scores=0.75850064,
final_local_scores=0.8329283, pem=0.0, similarity=28.945955]
9: 15573 Japan
    [pem_pos=0, final_scores=0.72510874,
final_local_scores=0.75201, pem=-0.13549899,
similarity=27.268055]
10: 15573 Japan
    [pem_pos=0, final_scores=0.72244745,
final_local_scores=0.7491029, pem=-0.13549899,
similarity=27.181301]
11: 3127711 Sandrine_Testud
    [pem_pos=0, final_scores=0.7143786,
final_local_scores=0.82386, pem=0.0, similarity=28.675339]
12: 5843419 France
    [pem_pos=0, final_scores=0.7141802,
final_local_scores=0.75052595, pem=-0.16240834,
similarity=27.370106]
13: 1514559 Ai_Sugiyama
    [pem_pos=0, final_scores=0.6942883,
final_local_scores=0.7702487, pem=0.0, similarity=27.07547]
14: 15573 Japan
    [pem_pos=0, final_scores=0.72001714,
final_local_scores=0.74644816, pem=-0.13549899,
similarity=27.10208]
15: 1028366 Nathalie_Tauziat
    [pem_pos=0, final_scores=0.79067117,
final_local_scores=0.9173444, pem=0.0, similarity=31.4651]
16: 5843419 France
    [pem_pos=0, final_scores=0.72030777,
final_local_scores=0.75721943, pem=-0.16240834,
similarity=27.569853]
17: 15340120 Wang_Shing-ting
    [pem_pos=0, final_scores=0.44481802,
final_local_scores=0.54129636, pem=0.0, similarity=20.243076]
18: 25734 Taiwan
    [pem_pos=0, final_scores=0.3929078,
final_local_scores=0.42178828, pem=-0.077886656,
similarity=17.100277]
```

Sample: b'35_TENNIS'

TENNIS - RESULTS AT HAMLET CUP . [1 1 COMMACK 1] , [2 2 New York 2] 1996-08-21 Results from the Waldbaum Hamlet Cup tennis tournament on Wednesday (prefix number denotes seeding) :
Second round 1 - [3 3 Michael Chang 3] ([4 4 U.S. 4]) beat [5 5 Sergi Bruguera 5] ([6 6 Spain 6]) 6-3 6-2 [7 7 Michael Joyce 7] ([8 8 U.S. 8]) beat 3 - [9 9 Richey Reneberg 9] ([10 10 U.S. 10]) 3-6 6-4 6-3 [11 11 Martin Damm 11] ([12 12 Czech Republic 12]) beat 6 - [13 13 Younes El Aynaoui 13] ([14

14 Morocco 14]) 5-7 6-3 3-0 retired [15 15 Karol Kucera 15]
([16 16 Slovakia 16]) beat [17 17 Hicham Arazi 17] ([18 18
Morocco 18]) 7-6 (7-4) 7-5

1: 127266 Commack,_New_York
2: 8210131 New_York
3: 391834 Michael_Chang
4: 3434750 United_States
5: 913625 Sergi_Bruguera
6: 26667 Spain
7: 4026170 Michael_Joyce_(tennis)
8: 3434750 United_States
9: 871728 Richey_Reneberg
10: 3434750 United_States
11: 4026614 Martin_Damm
12: 5321 Czech_Republic
13: 2090929 Younes_El_Aynaoui
14: 19291 Morocco
15: 3039956 Karol_Kučera
16: 26830 Slovakia
17: 2508144 Hicham_Arazi
18: 19291 Morocco

1: 127266 Commack,_New_York
[pem_pos=0, final_scores=0.49003357,
final_local_scores=0.6217844, pem=0.0, similarity=22.644999]
2: 8210131 New_York
[pem_pos=0, final_scores=0.17866836,
final_local_scores=0.18599263, pem=-0.2889045,
similarity=11.211211]
3: 391834 Michael_Chang
[pem_pos=0, final_scores=0.80170375,
final_local_scores=0.8634642, pem=0.0, similarity=29.857206]
4: 3434750 United_States
[pem_pos=0, final_scores=0.5069131,
final_local_scores=0.5058128, pem=-0.14075686,
similarity=19.949635]
5: 913625 Sergi_Bruguera
[pem_pos=0, final_scores=0.8411209,
final_local_scores=0.93966943, pem=0.0, similarity=32.131325]
6: 26667 Spain
[pem_pos=0, final_scores=0.6068514,
final_local_scores=0.6345727, pem=-0.18857051,
similarity=24.052103]
7: 4026170 Michael_Joyce_(tennis)
[pem_pos=1, final_scores=0.49414057,
final_local_scores=0.5934114, pem=-0.41227743,
similarity=24.040321]
8: 3434750 United_States
[pem_pos=0, final_scores=0.45958418,
final_local_scores=0.4541132, pem=-0.14075686,
similarity=18.406813]
9: 871728 Richey_Reneberg
[pem_pos=0, final_scores=0.69958895,

```
final_local_scores=0.8182308, pem=0.0, similarity=28.507353]
10: 3434750 United_States
    [pem_pos=0, final_scores=0.52126724,
final_local_scores=0.52149254, pem=-0.14075686,
similarity=20.417547]
11: 4026614 Martin_Damm
    [pem_pos=0, final_scores=0.8156133,
final_local_scores=0.93173736, pem=0.0, similarity=31.894613]
12: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.5575769,
final_local_scores=0.5751199, pem=-0.27794555,
similarity=22.76395]
13: 2090929 Younes_El_Aynaoui
    [pem_pos=0, final_scores=0.61664945,
final_local_scores=0.69785184, pem=0.0, similarity=24.915003]
14: 19291 Morocco
    [pem_pos=0, final_scores=0.5497853,
final_local_scores=0.5854931, pem=-0.29598162,
similarity=23.171589]
15: 3039956 Karol_Kučera
    [pem_pos=0, final_scores=0.7898217,
final_local_scores=0.8947124, pem=0.0, similarity=30.789717]
16: 26830 Slovakia
    [pem_pos=0, final_scores=0.6699042,
final_local_scores=0.69613767, pem=-0.36602598,
similarity=26.854357]
17: 2508144 Hicham_Arazi
    [pem_pos=0, final_scores=0.6279764,
final_local_scores=0.7374201, pem=0.0, similarity=26.095798]
18: 19291 Morocco
    [pem_pos=0, final_scores=0.60831845,
final_local_scores=0.64943165, pem=-0.29598162,
similarity=25.079641]
```

Sample: b'36_SOCCER'

SOCCKER - [1 1 DALGLISH 1] SAD OVER [2 2 BLACKBURN 2]
PARTING . [3 3 LONDON 3] 1996-08-22 [4 4 Kenny Dalglish 4]
spoke on Thursday of his sadness at leaving [5 5 Blackburn 5]
, the club he led to the [6 6 English 6] premier league title
in 1994-95 . [7 7 Blackburn 7] announced on Wednesday they
and [8 8 Dalglish 8] had parted by mutual consent . But the
ex-manager confessed on Thursday to being " sad " at leaving
after taking [9 9 Blackburn 9] from the second division to
the premier league title inside three and a half years . In a
telephone call to a local newspaper from his holiday home in [
10 10 Spain 10] , [11 11 Dalglish 11] said : " We came to the
same opinion , albeit the club came to it a little bit
earlier than me . " He added : " If no one asked , I never
opened my mouth . I have stayed out of the way and let them
get on with the job . The club thought it (the job) had run
its course and I came to the same conclusion . " [12 12 Dalg
ish 12] had been with [13 13 Blackburn 13] for nearly five

years , first as manager and then , for the past 15 months , as director of football .

1: 16919 Kenny_Dalglish
2: 298599 Blackburn_Rovers_F.C.
3: 17867 London
4: 16919 Kenny_Dalglish
5: 298599 Blackburn_Rovers_F.C.
6: 9316 England
7: 298599 Blackburn_Rovers_F.C.
8: 16919 Kenny_Dalglish
9: 298599 Blackburn_Rovers_F.C.
10: 26667 Spain
11: 16919 Kenny_Dalglish
12: 16919 Kenny_Dalglish
13: 298599 Blackburn_Rovers_F.C.
1: 16919 Kenny_Dalglish
[pem_pos=0, final_scores=0.65724736,
final_local_scores=0.6284119, pem=0.0, similarity=22.842777]
2: 298599 Blackburn_Rovers_F.C.
[pem_pos=1, final_scores=0.46099067,
final_local_scores=0.40069765, pem=-2.734066,
similarity=30.915607]
3: 17867 London
[pem_pos=0, final_scores=0.56114227,
final_local_scores=0.62542033, pem=-0.09086737,
similarity=23.247654]
4: 16919 Kenny_Dalglish
[pem_pos=0, final_scores=0.60084385,
final_local_scores=0.5667996, pem=0.0, similarity=21.004143]
5: 298599 Blackburn_Rovers_F.C.
[pem_pos=1, final_scores=0.5976392,
final_local_scores=0.54996544, pem=-2.734066,
similarity=35.370056]
6: 9316 England
[pem_pos=1, final_scores=0.6598034,
final_local_scores=0.7326123, pem=-1.3527626,
similarity=33.308857]
7: 298599 Blackburn_Rovers_F.C.
[pem_pos=1, final_scores=0.6041478,
final_local_scores=0.5570751, pem=-2.734066,
similarity=35.582222]
8: 16919 Kenny_Dalglish
[pem_pos=0, final_scores=0.6393918,
final_local_scores=0.6089074, pem=0.0, similarity=22.260725]
9: 298599 Blackburn_Rovers_F.C.
[pem_pos=1, final_scores=0.6697858,
final_local_scores=0.6287746, pem=-2.734066,
similarity=37.721878]
10: 26667 Spain
[pem_pos=0, final_scores=0.47648543,
final_local_scores=0.58703554, pem=-0.18857051,
similarity=22.6335]

```
11: 16919 Kenny_Dalglish
    [pem_pos=0, final_scores=0.49044997,
final_local_scores=0.44621104, pem=0.0, similarity=17.40554]
12: 16919 Kenny_Dalglish
    [pem_pos=0, final_scores=0.60978323,
final_local_scores=0.57656455, pem=0.0, similarity=21.29555]
13: 298599 Blackburn_Rovers_F.C.
    [pem_pos=1, final_scores=0.5893633,
final_local_scores=0.5409252, pem=-2.734066,
similarity=35.100277]
```

Sample: b'37_CRICKET'

CRICKET - [1 1 ENGLISH COUNTY CHAMPIONSHIP 1] SCORES . [2 2 L
ONDON 2] 1996-08-22 Close of play scores in four-day [1 1 Eng
lish 2 County Championship 1] cricket matches on Thursday :
Second day At [2 3 Weston-super-Mare 2] : [3 3 Durham 3] 326
(D. Cox 95 not out , S. Campbell 69 ; G. Rose 7-73) . [4 4 S
omerset 4] 236-4 (M. Lathwell 85) . First day At [5 5 Colch
ester 5] : [6 6 Gloucestershire 6] 280 ([7 7 J. Russell 7]
63 , A. Symonds 52 ; A. Cowan 5-68) . [8 8 Essex 8] 72-0 .
At [9 9 Cardiff 9] : [10 10 Kent 10] 128-1 (M. Walker 59 ,
D. Fulton 53 not out) v [11 11 Glamorgan 11] . At [12 12 Lei
cester 12] : [13 13 Leicestershire 13] 343-8 (P. Simmons
108 , P. Nixon 67 not out) v [14 14 Hampshire 14] . At [15 15
Northampton 15] : [16 16 Sussex 16] 368-7 (N. Lenham 145 ,
V. Drakes 59 not out , A. Wells 51) v [17 17 Northamptonshire
17] . At [18 18 Trent Bridge 18] : [19 19 Nottinghamshire 19]
392-6 (G. Archer 143 not out , M. Dowman 107) v [20 20 Surr
ey 20] . At [21 21 Worcester 21] : [22 22 Warwickshire 22]
255-9 (A. Giles 57 not out , W. Khan 52) v [23 23 Worcester
shire 23] . At [24 24 Headingley 24] : [25 25 Yorkshire 25]
305-5 (C. White 66 not out , M. Moxon 66 , M. Vaughan 57) v
[26 26 Lancashire 26] .

1: 683599 County_Championship

2: 17867 London

3: 1176371 Durham_County_Cricket_Club

4: 1622178 Somerset_County_Cricket_Club

5: 11568817 Castle_Park_Cricket_Ground

6: 1622322 Gloucestershire_County_Cricket_Club

7: 18417732 J._Russell_(Essex_cricketer)

8: 1622252 Essex_County_Cricket_Club

9: 5882 Cardiff

10: 1177719 Kent_County_Cricket_Club

11: 1622320 Glamorgan_County_Cricket_Club

12: 51155 Leicester

13: 1622318 Leicestershire_County_Cricket_Club

14: 1919370 Hampshire_County_Cricket_Club

15: 63734 Northampton

16: 1763691 Sussex_County_Cricket_Club

17: 530308 Northamptonshire_County_Cricket_Club

18: 849500 Trent_Bridge

19: 667341 Nottinghamshire_County_Cricket_Club

```
20: 766575 Surrey_County_Cricket_Club
21: 58681 Worcester
22: 667292 Warwickshire_County_Cricket_Club
23: 1156428 Worcestershire_County_Cricket_Club
24: 1876060 Headingley_Stadium
25: 869945 Yorkshire_County_Cricket_Club
26: 749919 Lancashire_County_Cricket_Club
1: 9316 England
    [winner: 683599 County_Championship, pem_pos=0,
final_scores=0.25344425, final_local_scores=0.18973213,
pem=0.0, similarity=9.7516985]
    RECALL MISS: 9316 England]
2: 683599 County_Championship
    [winner: 683599 County_Championship, pem_pos=0,
final_scores=0.25344425, final_local_scores=0.18973213,
pem=0.0, similarity=9.7516985]
    gt: 683599 County_Championship, pem_pos=0,
final_scores=0.25344425, final_local_scores=0.18973213,
pem=0.0, similarity=9.7516985 ]
3: 24597350 Clarence_Park,_Weston-super-Mare
    [winner: 330072 Weston-super-Mare, pem_pos=0,
final_scores=0.22668736, final_local_scores=0.232746,
pem=0.0, similarity=11.035318]
    gt: 24597350 Clarence_Park,_Weston-super-Mare,
pem_pos=4, final_scores=0.057522163,
final_local_scores=0.068338245, pem=-1.7623314,
similarity=15.712895 ]
1: 683599 County_Championship
    [pem_pos=0, final_scores=0.37056834,
final_local_scores=0.3176724, pem=0.0, similarity=13.569691]
2: 17867 London
    [pem_pos=0, final_scores=0.6814292,
final_local_scores=0.7426724, pem=-0.09086737,
similarity=26.746689]
3: 1176371 Durham_County_Cricket_Club
    [pem_pos=2, final_scores=0.84585565,
final_local_scores=0.8388098, pem=-2.3952909,
similarity=42.147434]
4: 1622178 Somerset_County_Cricket_Club
    [pem_pos=1, final_scores=0.6678099,
final_local_scores=0.626755, pem=-2.4139037,
similarity=35.920517]
5: 11568817 Castle_Park_Cricket_Ground
    [pem_pos=17, final_scores=0.25686014,
final_local_scores=0.31648886, pem=-2.976298,
similarity=29.719944]
6: 1622322 Gloucestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.7141813,
final_local_scores=0.67552394, pem=-1.379793,
similarity=31.752224]
7: 18417732 J._Russell_(Essex_cricketer)
    [pem_pos=7, final_scores=0.2597785,
```

```
final_local_scores=0.37836617, pem=-1.7917595,
similarity=25.12478]
8: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.83388215,
final_local_scores=0.8108008, pem=-2.3459642,
similarity=41.043343]
9: 5882 Cardiff
    [pem_pos=0, final_scores=0.65464395,
final_local_scores=0.6820928, pem=-0.20401932,
similarity=25.55421]
10: 1177719 Kent_County_Cricket_Club
    [pem_pos=1, final_scores=0.78736645,
final_local_scores=0.7572605, pem=-2.5096242,
similarity=40.335606]
11: 1622320 Glamorgan_County_Cricket_Club
    [pem_pos=1, final_scores=0.74077016,
final_local_scores=0.7082534, pem=-0.7903395,
similarity=29.523396]
12: 51155 Leicester
    [pem_pos=0, final_scores=0.598116,
final_local_scores=0.5963186, pem=-0.2625849,
similarity=23.313026]
13: 1622318 Leicestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.80325896,
final_local_scores=0.7807525, pem=-1.0465667,
similarity=33.08032]
14: 1919370 Hampshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.7563539,
final_local_scores=0.72939736, pem=-1.9232318,
similarity=36.315216]
15: 63734 Northampton
    [pem_pos=0, final_scores=0.5460656,
final_local_scores=0.5451284, pem=-0.5110059,
similarity=23.13636]
16: 1763691 Sussex_County_Cricket_Club
    [pem_pos=1, final_scores=0.68540055,
final_local_scores=0.657587, pem=-1.5566913,
similarity=32.17895]
17: 530308 Northamptonshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.40921843,
final_local_scores=0.34871453, pem=-1.2069724,
similarity=21.059753]
18: 849500 Trent_Bridge
    [pem_pos=0, final_scores=0.60922176,
final_local_scores=0.5952825, pem=0.0, similarity=21.85413]
19: 667341 Nottinghamshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6442383,
final_local_scores=0.5990398, pem=-1.1067253,
similarity=27.9848]
20: 766575 Surrey_County_Cricket_Club
    [pem_pos=1, final_scores=0.73441035,
final_local_scores=0.6957775, pem=-1.7952642,
```

```

similarity=34.616028]
21: 58681 Worcester
    [pem_pos=0, final_scores=0.53912085,
final_local_scores=0.55471015, pem=-0.59997296,
similarity=23.906116]
22: 667292 Warwickshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.7301194,
final_local_scores=0.69495535, pem=-1.1093602,
similarity=30.861439]
23: 1156428 Worcestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.70413584,
final_local_scores=0.6691755, pem=-0.9164197,
similarity=29.042877]
24: 1876060 Headingley_Stadium
    [pem_pos=0, final_scores=0.21669064,
final_local_scores=0.19300541, pem=-0.30334172,
similarity=11.498999]
25: 869945 Yorkshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6822546,
final_local_scores=0.6401022, pem=-1.7779728,
similarity=32.86053]
26: 749919 Lancashire_County_Cricket_Club
    [pem_pos=1, final_scores=0.45763797,
final_local_scores=0.3965224, pem=-1.7295884,
similarity=25.328503]
1: 683599 County_Championship
    [pem_pos=0, final_scores=0.25344425,
final_local_scores=0.18973213, pem=0.0, similarity=9.7516985]
2: 330072 Weston-super-Mare
    [pem_pos=0, final_scores=0.22668736,
final_local_scores=0.232746, pem=0.0, similarity=11.035318]

```

Sample: b'38_CRICKET'

CRICKET - [1 1 ENGLAND 1] V [2 2 PAKISTAN 2] FINAL TEST SCOREBOARD . [3 3 LONDON 3] 1996-08-22 Scoreboard on the first day of the third and final test between [4 4 England 4] and [5 5 Pakistan 5] at [6 6 The Oval 6] on Thursday : [7 7 England 7] first innings M. Atherton b [8 8 Waqar Younis 8] 31 A. Stewart b [9 9 Mushtaq Ahmed 9] 44 N. Hussain c [10 10 Saeed Anwar 10] b [11 11 Waqar Younis 11] 12 G. Thorpe lbw b [12 12 Mohammad Akram 12] 54 J. Crawley not out 94 N. Knight b [13 13 Mushtaq Ahmed 13] 17 C. Lewis b [14 14 Wasim Akram 14] 5 I. Salisbury not out 1 Extras (lb-11 w-1 nb-8) 20 Total (for six wickets) 278 Fall of wickets : 1-64 2-85 3-116 4-205 5-248 6-273 To bat : R. Croft , D. Cork , A. Mullally Bowling (to date) : [15 15 Wasim Akram 15] 25-8-61-1 , [16 16 Waqar Younis 16] 20-6-70-2 , [17 17 Mohammad Akram 17] 12-1-41-1 , [18 18 Mushtaq Ahmed 18] 27-5-78-2 , Aamir Sohail 6-1-17-0 [19 19 Pakistan 19] : Aamir Sohail , [20 20 Saeed Anwar 20] , [21 21 Ijaz Ahmed 21] , [22 22 Inzamam-ul-Haq 22] , [23 23 Salim Malik 23] , [24 24 Asif Mujtaba 24] , [25 25 Wasim Akram 25] , [26 26 Moin Khan 26] , [27 27 Mushtaq Ahmed 27]

```
, [28 28 Waqar Younis 28] , Mohammad Akam
1: 759125 England_cricket_team
2: 759295 Pakistan_national_cricket_team
3: 17867 London
4: 759125 England_cricket_team
5: 759295 Pakistan_national_cricket_team
6: 387687 The_Oval
7: 759125 England_cricket_team
8: 894433 Waqar_Younis
9: 896405 Mushtaq_Ahmed
10: 894437 Saeed_Anwar
11: 894433 Waqar_Younis
12: 1958830 Mohammad_Akram
13: 896405 Mushtaq_Ahmed
14: 501393 Wasim_Akram
15: 501393 Wasim_Akram
16: 894433 Waqar_Younis
17: 1958830 Mohammad_Akram
18: 896405 Mushtaq_Ahmed
19: 759295 Pakistan_national_cricket_team
20: 894437 Saeed_Anwar
21: 938863 Ijaz_Ahmed_(cricketer)
22: 896403 Inzamam-ul-Haq
23: 938861 Saleem_Malik
24: 3976606 Asif_Mujtaba
25: 501393 Wasim_Akram
26: 554132 Moin_Khan
27: 896405 Mushtaq_Ahmed
28: 894433 Waqar_Younis
1: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.625743,
final_local_scores=0.6006475, pem=-3.3054674,
similarity=39.98988]
2: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7271963,
final_local_scores=0.7023364, pem=-2.7556365,
similarity=40.03441]
3: 17867 London
    [pem_pos=0, final_scores=0.5114487,
final_local_scores=0.63929415, pem=-0.09086737,
similarity=23.661676]
4: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.7950074,
final_local_scores=0.78554314, pem=-3.3054674,
similarity=45.507534]
5: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.8507764,
final_local_scores=0.83732885, pem=-2.7556365,
similarity=44.062855]
6: 387687 The_Oval
    [pem_pos=0, final_scores=0.61156577,
final_local_scores=0.62176454, pem=0.0, similarity=22.644407]
```

7: 759125 England_cricket_team
[pem_pos=2, final_scores=0.75160426,
final_local_scores=0.7381318, pem=-3.3054674,
similarity=44.092686]
8: 894433 Waqar_Younis
[pem_pos=0, final_scores=0.706736,
final_local_scores=0.6890464, pem=0.0, similarity=24.652231]
9: 896405 Mushtaq_Ahmed
[pem_pos=0, final_scores=0.5732804,
final_local_scores=0.5536693, pem=0.0, similarity=20.612309]
10: 894437 Saeed_Anwar
[pem_pos=0, final_scores=0.7849172,
final_local_scores=0.78631556, pem=0.0, similarity=27.554937]
11: 894433 Waqar_Younis
[pem_pos=0, final_scores=0.65594006,
final_local_scores=0.63355947, pem=0.0, similarity=22.996391]
12: 1958830 Mohammad_Akram
[pem_pos=0, final_scores=0.52128166,
final_local_scores=0.53512496, pem=0.0, similarity=20.058908]
13: 896405 Mushtaq_Ahmed
[pem_pos=0, final_scores=0.60792726,
final_local_scores=0.5915157, pem=0.0, similarity=21.741722]
14: 501393 Wasim_Akram
[pem_pos=0, final_scores=0.71632344,
final_local_scores=0.69812983, pem=0.0, similarity=24.9233]
15: 501393 Wasim_Akram
[pem_pos=0, final_scores=0.75639564,
final_local_scores=0.7419027, pem=0.0, similarity=26.22957]
16: 894433 Waqar_Younis
[pem_pos=0, final_scores=0.7656775,
final_local_scores=0.753431, pem=0.0, similarity=26.573597]
17: 1958830 Mohammad_Akram
[pem_pos=0, final_scores=0.58053905,
final_local_scores=0.5998546, pem=0.0, similarity=21.99057]
18: 896405 Mushtaq_Ahmed
[pem_pos=0, final_scores=0.58239233,
final_local_scores=0.56362265, pem=0.0, similarity=20.909338]
19: 759295 Pakistan_national_cricket_team
[pem_pos=1, final_scores=0.7825298,
final_local_scores=0.7627798, pem=-2.7556365,
similarity=41.83816]
20: 894437 Saeed_Anwar
[pem_pos=0, final_scores=0.799607,
final_local_scores=0.8023619, pem=0.0, similarity=28.033792]
21: 938863 Ijaz_Ahmed_(cricketer)
[pem_pos=0, final_scores=0.66664857,
final_local_scores=0.6634691, pem=0.0, similarity=23.888954]
22: 896403 Inzamam-ul-Haq
[pem_pos=0, final_scores=0.980586,
final_local_scores=0.99212545, pem=0.0, similarity=33.696716]
23: 938861 Saleem_Malik
[pem_pos=0, final_scores=0.7096547,

```
final_local_scores=0.7300999, pem=0.0, similarity=25.87735]
24: 3976606 Asif_Mujtaba
    [pem_pos=0, final_scores=0.7346478,
final_local_scores=0.77165276, pem=0.0, similarity=27.11737]
25: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.77872235,
final_local_scores=0.7662912, pem=0.0, similarity=26.95737]
26: 554132 Moin_Khan
    [pem_pos=0, final_scores=0.7567268,
final_local_scores=0.7952809, pem=0.0, similarity=27.822481]
27: 896405 Mushtaq_Ahmed
    [pem_pos=0, final_scores=0.6260257,
final_local_scores=0.61128545, pem=0.0, similarity=22.33169]
28: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.7197919,
final_local_scores=0.70330787, pem=0.0, similarity=25.077822]
```

Sample: b'39_SOCCER'

SOCCER - [1 1 FERGUSON 1] BACK IN [2 2 SCOTTISH 2] SQUAD
AFTER 20 MONTHS . [3 3 GLASGOW 3] 1996-08-22 [4 4 Everton 4]
's [5 5 Duncan Ferguson 5] , who scored twice against [6 6 Ma
nchester United 6] on Wednesday , was picked on Thursday for
the [7 7 Scottish 7] squad after a 20-month exile . [8 8 Glas
gow Rangers 8] striker [9 9 Ally McCoist 9] , another man in
form after two hat-tricks in four days , was also named for
the August 31 [10 10 World Cup 10] qualifier against [11 11 Au
stria 11] in [12 12 Vienna 12] . [13 13 Ferguson 13] , who
served six weeks in jail in late 1995 for head-butting an
opponent , won the last of his five [14 14 Scotland 14] caps
in December 1994 . [15 15 Scotland 15] manager [16 16 Craig Br
own 16] said on Thursday : " I 've watched [17 17 Duncan Ferg
uson 17] in action twice recently and he 's bang in form . [18
18 Ally McCoist 18] is also in great scoring form at the
moment . " [19 19 Celtic 19] 's [20 20 Jackie McNamara 20] ,
who did well with last season 's successful under-21 team ,
earns a call-up to the senior squad .

```
1: 848405 Duncan_Ferguson
2: 578101 Scotland_national_football_team
3: 68736 Glasgow
4: 91155 Everton_F.C.
5: 848405 Duncan_Ferguson
6: 19961 Manchester_United_F.C.
7: 26994 Scotland
8: 69713 Rangers_F.C.
9: 838461 Ally_Mccoist
10: 126460 1998_FIFA_World_Cup
11: 695226 Austria_national_football_team
12: 55866 Vienna
13: 848405 Duncan_Ferguson
14: 578101 Scotland_national_football_team
15: 578101 Scotland_national_football_team
16: 883369 Craig_Brown_(footballer_born_1940)
```

```
17: 848405 Duncan_Ferguson
18: 838461 Ally_McCoist
19: 69708 Celtic_F.C.
20: 1585914 Jackie_McNamara
1: 848405 Duncan_Ferguson
    [pem_pos=0, final_scores=0.82940286,
final_local_scores=0.8352024, pem=0.0, similarity=29.013819]
2: 578101 Scotland_national_football_team
    [pem_pos=7, final_scores=0.37615126,
final_local_scores=0.32039773, pem=-3.9263487,
similarity=35.003117]
3: 68736 Glasgow
    [pem_pos=0, final_scores=0.47125202,
final_local_scores=0.5052313, pem=-0.15604196,
similarity=20.015404]
4: 91155 Everton_F.C.
    [pem_pos=0, final_scores=0.5617147,
final_local_scores=0.52005595, pem=-0.17171584,
similarity=20.543037]
5: 848405 Duncan_Ferguson
    [pem_pos=0, final_scores=0.64811176,
final_local_scores=0.63128114, pem=0.0, similarity=22.928402]
6: 19961 Manchester_United_F.C.
    [pem_pos=0, final_scores=0.4785288,
final_local_scores=0.43713582, pem=0.0, similarity=17.134718]
7: 26994 Scotland
    [pem_pos=0, final_scores=0.45596778,
final_local_scores=0.5076386, pem=-0.47951394,
similarity=21.846333]
8: 69713 Rangers_F.C.
    [pem_pos=0, final_scores=0.5616926,
final_local_scores=0.5206562, pem=0.0, similarity=19.627132]
9: 838461 Ally_McCoist
    [pem_pos=0, final_scores=0.6999759,
final_local_scores=0.70088595, pem=0.0, similarity=25.005548]
10: 126460 1998_FIFA_World_Cup
    [pem_pos=18, final_scores=0.3189094,
final_local_scores=0.3196491, pem=-4.2172847,
similarity=36.56293]
11: 695226 Austria_national_football_team
    [pem_pos=1, final_scores=0.45725906,
final_local_scores=0.46199852, pem=-2.931263,
similarity=33.817333]
12: 55866 Vienna
    [pem_pos=0, final_scores=0.39269245,
final_local_scores=0.4973442, pem=-0.12517071,
similarity=19.612152]
13: 848405 Duncan_Ferguson
    [pem_pos=0, final_scores=1.0174251,
final_local_scores=1.0399518, pem=0.0, similarity=35.12395]
14: 578101 Scotland_national_football_team
    [pem_pos=1, final_scores=0.2885934,
```

```
final_local_scores=0.22433943, pem=-2.6956348,
similarity=25.443731]
15: 578101 Scotland_national_football_team
    [pem_pos=1, final_scores=0.36851102,
final_local_scores=0.31163728, pem=-2.6956348,
similarity=28.048872]
16: 883369 Craig_Brown_(footballer_born_1940)
    [pem_pos=3, final_scores=0.4901836,
final_local_scores=0.47775167, pem=-1.974081,
similarity=29.082136]
17: 848405 Duncan_Ferguson
    [pem_pos=0, final_scores=0.61255264,
final_local_scores=0.59243816, pem=0.0, similarity=21.769249]
18: 838461 Ally_McCoist
    [pem_pos=0, final_scores=0.60678643,
final_local_scores=0.5990906, pem=0.0, similarity=21.967772]
19: 69708 Celtic_F.C.
    [pem_pos=0, final_scores=0.704415,
final_local_scores=0.67818743, pem=-0.9785444,
similarity=29.649656]
20: 1585914 Jackie_McNamara
    [pem_pos=0, final_scores=0.46101242,
final_local_scores=0.46568298, pem=0.0, similarity=17.986622]
```

Sample: b'40_CRICKET'

CRICKET - [1 1 ENGLAND 1] 100-2 AT LUNCH ON FIRST DAY OF
THIRD TEST . [2 2 LONDON 2] 1996-08-22 [3 3 England 3] were
100 for two at lunch on the first day of the third and final
test against [4 4 Pakistan 4] at [5 5 The Oval 5] on
Thursday .

```
1: 9316 England
2: 17867 London
3: 759125 England_cricket_team
4: 759295 Pakistan_national_cricket_team
5: 387687 The_Oval
1: 9316 England
    [pem_pos=0, final_scores=0.46998036,
final_local_scores=0.48060054, pem=-0.19053856,
similarity=19.46797]
2: 17867 London
    [pem_pos=0, final_scores=0.6495326,
final_local_scores=0.6905474, pem=-0.09086737,
similarity=25.191175]
3: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.6542182,
final_local_scores=0.63834006, pem=-3.3054674,
similarity=41.1147]
4: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.79296523,
final_local_scores=0.8285664, pem=-2.7556365,
similarity=43.801365]
5: 387687 The_Oval
```

```
[pem_pos=0, final_scores=0.5358327,  
final_local_scores=0.52771133, pem=0.0, similarity=19.837671]
```

Sample: b'41_SOCCER'

SOCCKER - [1 1 KEANE 1] SIGNS FOUR-YEAR CONTRACT WITH [2 2 MA
NCHESTER UNITED 2] . [3 3 LONDON 3] 1996-08-22 [4 4 Ireland 4]
midfielder [5 5 Roy Keane 5] has signed a new four-year
contract with [6 6 English 6] league and [7 7 F.A. Cup 7]
champions [8 8 Manchester United 8] . " 1 Roy agreed a new
deal before last night 's game against [9 9 Everton 9] and we
are delighted , " said [10 10 United 10] manager [11 11 Alex F
erguson 11] on Thursday .

1: 413261 Roy_Keane

2: 19961 Manchester_United_F.C.

3: 17867 London

4: 14560 Republic_of_Ireland

5: 413261 Roy_Keane

6: 9316 England

7: 11237 FA_Cup

8: 19961 Manchester_United_F.C.

9: 91155 Everton_F.C.

10: 19961 Manchester_United_F.C.

11: 43223 Alex_Ferguson

1: 413261 Roy_Keane

[winner: 3112042 Bryan_Roy, pem_pos=9,
final_scores=-0.23065653, final_local_scores=-0.23260146,
pem=-3.724678, similarity=17.4038]

RECALL MISS: 413261 Roy_Keane]

1: 413261 Roy_Keane

[pem_pos=0, final_scores=0.9008576,
final_local_scores=0.9204417, pem=0.0, similarity=31.55753]

2: 19961 Manchester_United_F.C.

[pem_pos=0, final_scores=0.57529694,
final_local_scores=0.5198512, pem=0.0, similarity=19.60311]

3: 17867 London

[pem_pos=0, final_scores=0.57281816,
final_local_scores=0.63544804, pem=-0.09086737,
similarity=23.5469]

4: 14560 Republic_of_Ireland

[pem_pos=1, final_scores=0.3678637,
final_local_scores=0.44494003, pem=-1.5040284,
similarity=25.546751]

5: 413261 Roy_Keane

[pem_pos=0, final_scores=0.66218764,
final_local_scores=0.6463945, pem=0.0, similarity=23.379414]

6: 9316 England

[pem_pos=1, final_scores=0.5296,
final_local_scores=0.5880708, pem=-1.3527626,
similarity=28.995451]

7: 11237 FA_Cup

[pem_pos=0, final_scores=0.48404676,
final_local_scores=0.4599889, pem=-0.16323638,

```

similarity=18.704403]
8: 19961 Manchester_United_F.C.
    [pem_pos=0, final_scores=0.4800037,
final_local_scores=0.41575783, pem=0.0, similarity=16.496756]
9: 91155 Everton_F.C.
    [pem_pos=0, final_scores=0.55940443,
final_local_scores=0.5073287, pem=-0.17171584,
similarity=20.16323]
10: 19961 Manchester_United_F.C.
    [pem_pos=5, final_scores=0.5408132,
final_local_scores=0.4835071, pem=-3.3400002,
similarity=36.68197]
11: 43223 Alex_Ferguson
    [pem_pos=0, final_scores=0.48059338,
final_local_scores=0.4551103, pem=0.0, similarity=17.671112]

```

Sample: b'42_TENNIS'

TENNIS - RESULTS AT [1 1 CANADIAN OPEN 1] . [2 2 TORONTO 2]
1996-08-21 Results from the [3 3 Canadian Open 3] tennis
tournament on Wednesday (prefix number denotes seeding) :
Second round [4 4 Daniel Nestor 4] ([5 5 Canada 5]) beat 1
- [6 6 Thomas Muster 6] ([7 7 Austria 7]) 6-3 7-5 [8 8 Mika
el Tillstrom 8] ([9 9 Sweden 9]) beat 2 - [10 10 Goran Ivan
isevic 10] ([11 11 Croatia 11]) 6-7 (3-7) 6-4 6-4 3 - [12
12 Wayne Ferreira 12] ([13 13 South Africa 13]) beat [14 14
Jiri Novak 14] ([15 15 Czech Republic 15]) 7-5 6-3 4 - [16 1
6 Marcelo Rios 16] ([17 17 Chile 17]) beat [18 18 Kenneth Ca
rlsen 18] ([19 19 Denmark 19]) 6-3 6-2 6 - [20 20 MaliVai Wa
shington 20] ([21 21 U.S. 21]) beat [22 22 Alex Corretja 22]
([23 23 Spain 23]) 6-4 6-2 7 - [24 24 Todd Martin 24] ([25
25 U.S. 25]) beat [26 26 Renzo Furlan 26] ([27 27 Italy 27]
) 7-6 (7-3) 6-3 [28 28 Mark Philippoussis 28] ([29 29 Aust
ralia 29]) beat 8 - [30 30 Marc Rosset 30] ([31 31 Switzerl
and 31]) 6-3 3-6 7-6 (8-6) 9 - [32 32 Cedric Pioline 32] ([33
33 France 33]) beat [34 34 Gregory Carraz 34] ([35 35 Fr
ance 35]) 7-6 (7-1) 6-4 [36 36 Patrick Rafter 36] ([37 37
Australia 37]) beat 11 - [38 38 Alberto Berasategui 38] ([39
39 Spain 39]) 6-1 6-2 [40 40 Petr Korda 40] ([41 41 Czech Re
public 41]) beat 12 - [42 42 Francisco Clavet 42] ([43 43 Sp
ain 43]) 6-3 6-4 [44 44 Daniel Vacek 44] ([45 45 Czech Rep
ublic 45]) beat 13 - [46 46 Jason Stoltenberg 46] ([47 47 Au
stralia 47]) 5-7 7-6 (7-1) 7-6 (13-11) [48 48 Todd Woodb
ridge 48] ([49 49 Australia 49] beat [50 50 Sebastien Lareau
50] ([51 51 Canada 51]) 6-3 1-6 6-3 Alex O'Brien ([52 52 U.
S. 52]) beat [53 53 Byron Black 53] ([54 54 Zimbabwe 54])
7-6 (7-2) 6-2 [55 55 Bohdan Ulihrach 55] ([56 56 Czech Rep
ublic 56]) beat [57 57 Andrea Gaudenzi 57] ([58 58 Italy 58]
) 6-3 4-6 6-1 [59 59 Tim Henman 59] ([60 60 Britain 60])
beat [61 61 Chris Woodruff 61] ([62 62 U.S. 62]) , walkover
1: 1872219 Canadian_Open_(tennis)
2: 64646 Toronto
3: 1872219 Canadian_Open_(tennis)

4: 998494 Daniel_Nestor
5: 5042916 Canada
6: 166170 Thomas_Muster
7: 26964606 Austria
8: 3943969 Mikael_Tillström
9: 5058739 Sweden
10: 194244 Goran_Ivanišević
11: 5573 Croatia
12: 1723516 Wayne_Ferreira
13: 17416221 South_Africa
14: 1554049 Jiří_Novák
15: 5321 Czech_Republic
16: 483554 Marcelo_Ríos
17: 5489 Chile
18: 1804967 Kenneth_Carlsen
19: 76972 Denmark
20: 873616 MaliVai_Washington
21: 3434750 United_States
22: 17733765 Àlex_Corretja
23: 26667 Spain
24: 871725 Todd_Martin
25: 3434750 United_States
26: 3929800 Renzo_Furlan
27: 14532 Italy
28: 261368 Mark_Philippoussis
29: 4689264 Australia
30: 1717291 Marc_Rosset
31: 26748 Switzerland
32: 979076 Cédric_Pioline
33: 5843419 France
34: 20083437 Grégory_Carraz
35: 5843419 France
36: 23297 Patrick_Rafter
37: 4689264 Australia
38: 1728009 Alberto_Berasategui
39: 26667 Spain
40: 1439769 Petr_Korda
41: 5321 Czech_Republic
42: 4076952 Francisco_Clavet
43: 26667 Spain
44: 3930665 Daniel_Vacek
45: 5321 Czech_Republic
46: 2598440 Jason_Stoltenberg
47: 4689264 Australia
48: 1547951 Todd_Woodbridge
49: 4689264 Australia
50: 1954361 Sébastien_Lareau
51: 5042916 Canada
52: 3434750 United_States
53: 1869714 Byron_Black
54: 34399 Zimbabwe
55: 2805841 Bohdan_Ulihrach

```
56: 5321 Czech_Republic
57: 3929637 Andrea_Gaudenzi
58: 14532 Italy
59: 262030 Tim_Henman
60: 31717 United_Kingdom
61: 2515285 Chris_Woodruff
62: 3434750 United_States
1: 1872219 Canadian_Open_(tennis)
    [pem_pos=0, final_scores=0.5222374,
final_local_scores=0.5564929, pem=-0.58982956,
similarity=23.904156]
2: 64646 Toronto
    [pem_pos=0, final_scores=0.5678237,
final_local_scores=0.6127704, pem=-0.12291032,
similarity=23.044409]
3: 1872219 Canadian_Open_(tennis)
    [pem_pos=0, final_scores=0.4375518,
final_local_scores=0.4639867, pem=-0.58982956,
similarity=21.143585]
4: 998494 Daniel_Nestor
    [pem_pos=0, final_scores=0.7323025,
final_local_scores=0.8067406, pem=0.0, similarity=28.164461]
5: 5042916 Canada
    [pem_pos=0, final_scores=0.70089245,
final_local_scores=0.73719764, pem=-0.11662793,
similarity=26.723402]
6: 166170 Thomas_Muster
    [pem_pos=0, final_scores=0.9444811,
final_local_scores=1.0073378, pem=0.0, similarity=34.15068]
7: 26964606 Austria
    [pem_pos=0, final_scores=0.6317594,
final_local_scores=0.6468905, pem=-0.25262418,
similarity=24.768026]
8: 3943969 Mikael_Tillström
    [pem_pos=0, final_scores=0.6677476,
final_local_scores=0.7622818, pem=0.0, similarity=26.83772]
9: 5058739 Sweden
    [pem_pos=0, final_scores=0.7316446,
final_local_scores=0.7647828, pem=-0.2135508,
similarity=28.073679]
10: 194244 Goran_Ivanišević
    [pem_pos=0, final_scores=0.97737926,
final_local_scores=1.0515525, pem=0.0, similarity=35.47014]
11: 5573 Croatia
    [pem_pos=0, final_scores=0.60688597,
final_local_scores=0.6302557, pem=-0.35021794,
similarity=24.80234]
12: 1723516 Wayne_Ferreira
    [pem_pos=0, final_scores=0.77525216,
final_local_scores=0.86444765, pem=0.0, similarity=29.886555]
13: 17416221 South_Africa
    [pem_pos=0, final_scores=0.5648341,
```

```
final_local_scores=0.6210476, pem=-0.26251382,
similarity=24.050604]
14: 1554049 Jiří_Novák
    [pem_pos=0, final_scores=0.87009686,
final_local_scores=0.96079177, pem=0.0, similarity=32.761654]
15: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.5865811,
final_local_scores=0.59744847, pem=-0.27794555,
similarity=23.430279]
16: 483554 Marcelo_Ríos
    [pem_pos=0, final_scores=0.634087,
final_local_scores=0.67058724, pem=0.0, similarity=24.101374]
17: 5489 Chile
    [pem_pos=0, final_scores=0.6380863,
final_local_scores=0.6878645, pem=-0.27404457,
similarity=26.107258]
18: 1804967 Kenneth_Carlsen
    [pem_pos=0, final_scores=0.65194947,
final_local_scores=0.72541595, pem=0.0, similarity=25.737572]
19: 76972 Denmark
    [pem_pos=0, final_scores=0.5734553,
final_local_scores=0.6075531, pem=-0.26762357,
similarity=23.675688]
20: 873616 MaliVai_Washington
    [pem_pos=0, final_scores=0.83924514,
final_local_scores=0.9102158, pem=0.0, similarity=31.252369]
21: 3434750 United_States
    [pem_pos=0, final_scores=0.4892366,
final_local_scores=0.50219446, pem=-0.14075686,
similarity=19.841654]
22: 17733765 Àlex_Corretja
    [pem_pos=0, final_scores=0.90786666,
final_local_scores=0.9740997, pem=0.0, similarity=33.15879]
23: 26667 Spain
    [pem_pos=0, final_scores=0.6458066,
final_local_scores=0.66903615, pem=-0.18857051,
similarity=25.080563]
24: 871725 Todd_Martin
    [pem_pos=0, final_scores=0.8243452,
final_local_scores=0.88923484, pem=0.0, similarity=30.626253]
25: 3434750 United_States
    [pem_pos=0, final_scores=0.48832995,
final_local_scores=0.5012041, pem=-0.14075686,
similarity=19.812101]
26: 3929800 Renzo_Furlan
    [pem_pos=0, final_scores=0.78069335,
final_local_scores=0.8867188, pem=0.0, similarity=30.55117]
27: 14532 Italy
    [pem_pos=0, final_scores=0.69757503,
final_local_scores=0.74952483, pem=-0.20612453,
similarity=27.577967]
28: 261368 Mark_Philippoussis
```

```
[pem_pos=0, final_scores=0.8358653,
final_local_scores=0.89437765, pem=0.0, similarity=30.779726]
29: 4689264 Australia
    [pem_pos=0, final_scores=0.6093077,
final_local_scores=0.64295423, pem=-0.17168477,
similarity=24.2104]
30: 1717291 Marc_Rosset
    [pem_pos=0, final_scores=0.6458686,
final_local_scores=0.72208345, pem=0.0, similarity=25.638123]
31: 26748 Switzerland
    [pem_pos=0, final_scores=0.648073,
final_local_scores=0.6608935, pem=-0.18598008,
similarity=24.82348]
32: 979076 Cédric_Pioline
    [pem_pos=0, final_scores=0.8399504,
final_local_scores=0.91407174, pem=0.0, similarity=31.367437]
33: 5843419 France
    [pem_pos=0, final_scores=0.6775198,
final_local_scores=0.69742763, pem=-0.16240834,
similarity=25.785545]
34: 20083437 Grégory_Carraz
    [pem_pos=0, final_scores=0.63717705,
final_local_scores=0.788491, pem=0.0, similarity=27.619858]
35: 5843419 France
    [pem_pos=0, final_scores=0.66797626,
final_local_scores=0.6870028, pem=-0.16240834,
similarity=25.474447]
36: 23297 Patrick_Rafter
    [pem_pos=0, final_scores=0.99395233,
final_local_scores=1.0726411, pem=0.0, similarity=36.09946]
37: 4689264 Australia
    [pem_pos=0, final_scores=0.6113289,
final_local_scores=0.64516217, pem=-0.17168477,
similarity=24.276287]
38: 1728009 Alberto_Berasategui
    [pem_pos=0, final_scores=0.65594995,
final_local_scores=0.7160574, pem=0.0, similarity=25.458294]
39: 26667 Spain
    [pem_pos=0, final_scores=0.5773079,
final_local_scores=0.5942117, pem=-0.18857051,
similarity=22.84765]
40: 1439769 Petr_Korda
    [pem_pos=0, final_scores=0.8741248,
final_local_scores=0.96299654, pem=0.0, similarity=32.82745]
41: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.56776124,
final_local_scores=0.5768906, pem=-0.27794555,
similarity=22.81679]
42: 4076952 Francisco_Clavet
    [pem_pos=0, final_scores=0.656711,
final_local_scores=0.7628842, pem=0.0, similarity=26.8557]
43: 26667 Spain
```

```
[pem_pos=0, final_scores=0.62253886,
final_local_scores=0.6436196, pem=-0.18857051,
similarity=24.32208]
44: 3930665 Daniel_Vacek
    [pem_pos=0, final_scores=0.6910089,
final_local_scores=0.7916826, pem=0.0, similarity=27.715101]
45: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.5524283,
final_local_scores=0.5601417, pem=-0.27794555,
similarity=22.316969]
46: 2598440 Jason_Stoltenberg
    [pem_pos=0, final_scores=0.7963782,
final_local_scores=0.8746665, pem=0.0, similarity=30.191507]
47: 4689264 Australia
    [pem_pos=0, final_scores=0.6634287,
final_local_scores=0.7020732, pem=-0.17168477,
similarity=25.974627]
48: 1547951 Todd_Woodbridge
    [pem_pos=0, final_scores=0.8192001,
final_local_scores=0.9046274, pem=0.0, similarity=31.0856]
49: 4689264 Australia
    [pem_pos=0, final_scores=0.7382329,
final_local_scores=0.78378546, pem=-0.17168477,
similarity=28.413084]
50: 1954361 Sébastien_Lareau
    [pem_pos=0, final_scores=0.5582256,
final_local_scores=0.63989305, pem=0.0, similarity=23.185398]
51: 5042916 Canada
    [pem_pos=0, final_scores=0.7222346,
final_local_scores=0.76051074, pem=-0.11662793,
similarity=27.419111]
52: 3434750 United_States
    [pem_pos=0, final_scores=0.4838819,
final_local_scores=0.49634522, pem=-0.14075686,
similarity=19.6671]
53: 1869714 Byron_Black
    [pem_pos=0, final_scores=0.71070826,
final_local_scores=0.8044049, pem=0.0, similarity=28.09476]
54: 34399 Zimbabwe
    [pem_pos=0, final_scores=0.569391,
final_local_scores=0.646177, pem=-0.36737105,
similarity=25.370745]
55: 2805841 Bohdan_Ulihrach
    [pem_pos=0, final_scores=0.8172603,
final_local_scores=0.9340623, pem=0.0, similarity=31.963997]
56: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.5937578,
final_local_scores=0.6052879, pem=-0.27794555,
similarity=23.664223]
57: 3929637 Andrea_Gaudenzi
    [pem_pos=0, final_scores=0.992214,
final_local_scores=1.1225712, pem=0.0, similarity=37.589478]
```

```
58: 14532 Italy
    [pem_pos=0, final_scores=0.7572014,
final_local_scores=0.81465757, pem=-0.20612453,
similarity=29.521658]
59: 262030 Tim_Henman
    [pem_pos=0, final_scores=0.85061246,
final_local_scores=0.893925, pem=0.0, similarity=30.766218]
60: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5257209,
final_local_scores=0.56040615, pem=-0.61114657,
similarity=24.13686]
61: 2515285 Chris_Woodruff
    [pem_pos=0, final_scores=0.7186075,
final_local_scores=0.8230921, pem=0.0, similarity=28.652424]
62: 3434750 United_States
    [pem_pos=0, final_scores=0.50086325,
final_local_scores=0.5148948, pem=-0.14075686,
similarity=20.220657]
```

Sample: b'43_CRICKET'

CRICKET - [1 1 MILLNS 1] SIGNS FOR [2 2 BOLAND 2] . [3 3 CAPE TOWN 3] 1996-08-22 [1 South African 1] provincial side [4 4 Boland 4] said on Thursday they had signed [5 5 Leicestershire 5] fast bowler [6 6 David Millns 6] on a one year contract . Millns , who toured [7 7 Australia 7] with [8 8 England 8] A in 1992/93 , replaces former [9 9 England 9] all-rounder [10 10 Phillip DeFreitas 10] as [11 11 Boland 11] 's overseas professional .

```
1: 10300919 David_Millns
2: 5570494 Boland_cricket_team
3: 6653 Cape_Town
4: 5570494 Boland_cricket_team
5: 1622318 Leicestershire_County_Cricket_Club
6: 10300919 David_Millns
7: 4689264 Australia
8: 9316 England
9: 759125 England_cricket_team
10: 2047474 Phillip_DeFreitas
11: 5570494 Boland_cricket_team
1: 10300919 David_Millns
    [pem_pos=0, final_scores=0.7316258,
final_local_scores=0.8161135, pem=0.0, similarity=28.444166]
2: 5570494 Boland_cricket_team
    [pem_pos=0, final_scores=0.49412906,
final_local_scores=0.49332255, pem=-0.52594304,
similarity=21.6716]
3: 6653 Cape_Town
    [pem_pos=0, final_scores=0.33423048,
final_local_scores=0.37579197, pem=0.0, similarity=15.304095]
4: 5570494 Boland_cricket_team
    [pem_pos=0, final_scores=0.4790905,
final_local_scores=0.4768952, pem=-0.52594304,
```

```

similarity=21.181377]
5: 1622318 Leicestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6741691,
final_local_scores=0.68729633, pem=-1.0465667,
similarity=30.291399]
6: 10300919 David_Millns
    [pem_pos=0, final_scores=0.49684852,
final_local_scores=0.53517556, pem=0.0, similarity=20.06042]
7: 4689264 Australia
    [pem_pos=0, final_scores=0.583305,
final_local_scores=0.67717445, pem=-0.17168477,
similarity=25.231598]
8: 9316 England
    [pem_pos=0, final_scores=0.51122576,
final_local_scores=0.56974125, pem=-0.19053856,
similarity=22.128107]
9: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.72407717,
final_local_scores=0.72322017, pem=-3.3054674,
similarity=43.64769]
10: 2047474 Phillip_DeFreitas
    [pem_pos=0, final_scores=0.5941358,
final_local_scores=0.6485282, pem=0.0, similarity=23.443089]
11: 5570494 Boland_cricket_team
    [pem_pos=0, final_scores=0.5065877,
final_local_scores=0.5069317, pem=-0.52594304,
similarity=22.077726]
1: 759263 South_Africa_national_cricket_team
    [pem_pos=1, final_scores=0.10049726,
final_local_scores=0.051888525, pem=-2.617051,
similarity=19.8701]

```

Sample: b'44_SOCCER'

SOCCER - EUROPEAN CUP WINNERS ' CUP RESULTS . [1 1 TIRANA 1]
1996-08-22 Results of European Cup Winners ' Cup qualifying
round , second leg soccer matches on Thursday : In [2 2 Tirana
2] : [3 3 Flamurtari Vlore 3] ([4 4 Albania 4]) 0 Chemlon
Humenne ([5 5 Slovakia 5]) 2 (halftime 0-0) Scorers :
Lubarskij (50th minute) , Valkucak (54th) Attendance :
5,000 Chemlon Humenne win 3-0 on aggregate In [6 6 Bistrita 6]
: [7 7 Gloria Bistrita 7] ([8 8 Romania 8]) 2 [9 9 Valletta
9] ([10 10 Malta 10]) 1 (1-1) Scorers : [11 11 Gloria Bis
trita 11] - Ilie Lazar (32nd) , Eugen Voica (84th) [12 12
Valletta 12] - [13 13 Gilbert Agius 13] (24th) Attendance :
8,000 [14 14 Gloria Bistrita 14] win 4-2 on aggregate . In [15
15 Chorzow 15] : [16 16 Ruch Chorzow 16] ([17 17 Poland
17]) 5 [18 18 Llansantffraid 18] ([19 19 Wales 19]) 0
(1-0) Scorers : [20 20 Arkadiusz Bak 20] (1st and 55th) ,
Arwel Jones (47th , own goal) , Miroslav Bak (62nd and
63rd) Attendance : 6,500 [21 21 Ruch Chorzow 21] win 6-1 on
aggregate In [22 22 Larnaca 22] : [23 23 AEK Larnaca 23] ([24
24 Cyprus 24]) 5 Kotaik Aboyan ([25 25 Armenia 25]) 0

(2-0) Scorers : Zoran Kundic (28th) , Klimis Alexandrou (41st) , Milenko Kovasevic (60th , penalty) , Goran Koprinovic (82nd) , Pavlos Markou (84th) Attendance : 5,000 [26 26 AEK Larnaca 26] win 5-1 on aggregate In [27 27 Si auliai 27] : [28 28 Kareda Siauliai 28] ([29 29 Lithuania 29]) 0 [30 30 Sion 30] ([31 31 Switzerland 31]) 0 Attendance : 5,000 [32 32 Sion 32] win 4-2 on aggregate . In [33 33 Vinnyt sya 33] : Nyva Vinnytsya ([34 34 Ukraine 34]) 1 [35 35 Tall inna Sadam 35] ([36 36 Estonia 36]) 0 (0-0) Attendance : 3,000 Aggregate score 2-2 . [37 37 Nyva 37] qualified on away goals rule . In [38 38 Bergen 38] : [39 39 Brann 39] ([40 40 Norway 40]) 2 [41 41 Shelbourne 41] ([42 42 Ireland 42]) 1 (1-1) Scorers : [43 43 Brann 43] - [44 44 Mons Ivar Mjelde 44] (10th) , [45 45 Jan Ove Pedersen 45] (72nd) [46 46 She lbourne 46] - [47 47 Mark Rutherford 47] (5th) Attendance : 2,189 [48 48 Brann 48] win 5-2 on aggregate In [49 49 Sofia 49] : [50 50 Levski Sofia 50] ([51 51 Bulgaria 51]) 1 [52 52 O limpija 52] ([53 53 Slovenia 53]) 0 (0-0) Scorer : Ilian Simeonov (58th) Attendance : 25,000 Aggregate 1-1 . [54 54 O limpija 54] won 4-3 on penalties . In [55 55 Vaduz 55] : [56 56 Vaduz 56] ([57 57 Liechtenstein 57]) 1 RAF Riga ([58 58 Latvia 58]) 1 (0-0) Scorers : [59 59 Vaduz 59] - Daniele Polverino (90th) RAF Riga - Agrins Zarins (47th) Aggregate 2-2 . [60 60 Vaduz 60] won 4-2 on penalties . In [61 61 Luxembourg 61] : [62 62 US Luxembourg 62] ([63 63 Luxembo urg 63]) 0 1 Varteks Varazdin ([64 64 Croatia 64]) 3 (0-0) Scorers : Drazen Beser (63rd) , Miljenko Mumler (penalty , 78th) , Jamir Cvetko (87th) Attendance : 800 2 Varteks Varazdin win 5-1 on aggregate . In [65 65 Torshavn 65] : [66 66 Havnar Boltfelag 66] ([67 67 Faroe Islands 67]) 0 [68 68 Dynamo Batumi 68] ([69 69 Georgia 69]) 3 (0-2) [70 70 Dynamo Batumi 70] win 9-0 on aggregate . In [71 71 Prague 71] : [72 72 Sparta Prague 72] ([73 73 Czech Republic 73]) 8 [74 74 Glentoran 74] ([75 75 Northern Ireland 75]) 0 (4-0) Scorers : Petr Gunda (1st and 26th) , Lumir Mistr (19th) , [76 76 Horst Siegl 76] (24th , 48th , 80th) , [77 77 Zdenek Svoboda 77] (76th) , [78 78 Petr Gabriel 78] (86th) 3 Sparta win 10-1 on aggregate . In [79 79 Edinburgh 79] : [80 80 Hearts 80] ([81 81 Scotland 81]) 1 [82 82 Red Star Belgrade 82] ([83 83 Yugoslavia 83]) 1 (1-0) Scorers : [84 84 Hearts 84] - [85 85 Dave McPherson 85] (44th) [86 86 Red Star 86] - [87 87 Vinko Marinovic 87] (59th) Attendance : 15,062 Aggregate 1-1 . [88 88 Red Star 88] win on away goals rule . In Rishon-Lezion : [89 89 Hapoel Ironi 89] ([90 90 Israel 90]) 3 [91 91 Constructorul Chisinau 91] ([92 92 Moldova 92]) 2 (2-1) Aggregate 3-3 . Constructorul win on away goals rule . In Anjalonkoski : [93 93 MyPa-47 93] ([94 94 Finland 94]) 1 Karabach Agdam ([95 95 Azerbaijan 95]) 1 (0-0) MyPa-47 win 2-1 on aggregate . In [96 96 Skopje 96] : 4 Sloga Jugomagnat ([97 97 Macedonia 97]) 0 [98 98 Kispest Honved 98] ([99 99 Hungary 99]) 1 (0-0) [100 100 Kispest Honved 100] win 2-0 on aggregate .

Add Hapoel Ironi v [101 101 Constructorul Chisinau 101]
Scorers : Rishon - Moshe Sabag (10th minute) , Nissan
Kapeta (26th) , Tomas Cibola (58th) . Constructorul -
Sergei Rogachev (42nd) , Gennadi Skidan (87th) .
Attendance : 1,500 .

1: 31075 Tirana
2: 31075 Tirana
3: 6573050 Flamurtari_Vlorë
4: 738 Albania
5: 26830 Slovakia
6: 415846 Bistrita
7: 2384775 ACF_Gloria_Bistrița
8: 25445 Romania
9: 2990151 Valletta_F.C.
10: 1009065 Malta_national_football_team
11: 2384775 ACF_Gloria_Bistrița
12: 2990151 Valletta_F.C.
13: 10424084 Gilbert_Agius
14: 2384775 ACF_Gloria_Bistrița
15: 74612 Chorzów
16: 677685 Ruch_Chorzów
17: 771860 Poland_national_football_team
18: 300136 The_New_Saints_F.C.
19: 580291 Wales_national_football_team
20: 2996295 Arkadiusz_Bąk
21: 677685 Ruch_Chorzów
22: 287852 Larnaca
23: 3550878 AEK_Larnaca_F.C.
24: 722717 Cyprus_national_football_team
25: 695292 Armenia_national_football_team
26: 3550878 AEK_Larnaca_F.C.
27: 404466 Šiauliai
28: 6492937 FK_Kareda_Kaunas
29: 17675 Lithuania
30: 1061075 FC_Sion
31: 681862 Switzerland_national_football_team
32: 1061075 FC_Sion
33: 1213985 Vinnytsia
34: 1008682 Ukraine_national_football_team
35: 6680744 JK_Tallinna_Sadam
36: 28222445 Estonia
37: 7010481 PFC_Nyva_Vinnytsia
38: 56494 Bergen
39: 192068 SK_Brann
40: 21241 Norway
41: 830612 Shelbourne_F.C.
42: 147575 Ireland
43: 192068 SK_Brann
44: 3118127 Mons_Ivar_Mjelde
45: 13926052 Jan_Ove_Pedersen
46: 830612 Shelbourne_F.C.
47: 6056077 Mark_Rutherford_(footballer)

48: 192068 SK_Brann
49: 57644 Sofia
50: 893469 PFC_Levski_Sofia
51: 3415 Bulgaria
52: 27792964 NK_Olimpija_Ljubljana_(defunct)
53: 27338 Slovenia
54: 27792964 NK_Olimpija_Ljubljana_(defunct)
55: 754336 Rheinpark_Stadion
56: 827884 FC_Vaduz
57: 652703 Liechtenstein_national_football_team
58: 17514 Latvia
59: 827884 FC_Vaduz
60: 827884 FC_Vaduz
61: 57035 Luxembourg_(city)
62: 5337201 Union_Luxembourg
63: 17515 Luxembourg
64: 715240 Croatia_national_football_team
65: 43979 Tórshavn
66: 831312 Havnar_Bóltfelag
67: 1011425 Faroe_Islands_national_football_team
68: 6581309 FC_Dinamo_Batumi
69: 48768 Georgia_(country)
70: 6581309 FC_Dinamo_Batumi
71: 23844 Prague
72: 861331 AC_Sparta_Prague
73: 729739 Czech_Republic_national_football_team
74: 828172 Glentoran_F.C.
75: 426388 Northern_Ireland_national_football_team
76: 5363309 Horst_Siegl
77: 12633520 Zdeněk_Svoboda
78: 5364238 Petr_Gabriel
79: 9602 Edinburgh
80: 302212 Heart_of_Midlothian_F.C.
81: 26994 Scotland
82: 361966 Red_Star_Belgrade
83: 34244 Yugoslavia
84: 302212 Heart_of_Midlothian_F.C.
85: 6789447 Dave_McPherson_(footballer)
86: 361966 Red_Star_Belgrade
87: 23671907 Vinko_Marinović
88: 361966 Red_Star_Belgrade
89: 5897396 Hapoel_Rishon_LeZion_F.C.
90: 9282173 Israel
91: 833887 FC_Tiraspol
92: 1011292 Moldova_national_football_team
93: 2079306 MYPA
94: 10577 Finland
95: 655906 Azerbaijan_national_football_team
96: 29618 Skopje
97: 23564616 Republic_of_Macedonia
98: 833965 Budapest_Honvéd_FC
99: 13275 Hungary

100: 833965 Budapest_Honvéd_FC
101: 833887 FC_Tiraspol
1: 1561948 NK_Varaždin_ŠN
[winner: 37739107 NK_Varaždin, pem_pos=0,
final_scores=-0.13684794, final_local_scores=-0.18637432,
pem=0.0, similarity=-1.4720684]
RECALL MISS: 1561948 NK_Varaždin_ŠN]
2: 1561948 NK_Varaždin_ŠN
[winner: 37739107 NK_Varaždin, pem_pos=0,
final_scores=-0.08792457, final_local_scores=-0.13293298,
pem=0.0, similarity=0.12272777]
RECALL MISS: 1561948 NK_Varaždin_ŠN]
3: 861331 AC_Sparta_Prague
[winner: 79818 Sparta_Rotterdam, pem_pos=2,
final_scores=-0.033357233, final_local_scores=-0.058967948,
pem=-2.9777987, similarity=18.523726]
RECALL MISS: 861331 AC_Sparta_Prague]
4: 828168 FK_Shkupi
[winner: 40070758 FK_Sloga_Jugomagnat, pem_pos=0,
final_scores=-0.20772213, final_local_scores=-0.24695957,
pem=0.0, similarity=-3.2800534]
RECALL MISS: 828168 FK_Shkupi]
1: 31075 Tirana
[pem_pos=0, final_scores=0.39379126,
final_local_scores=0.41647553, pem=0.0, similarity=16.518173]
2: 31075 Tirana
[pem_pos=0, final_scores=0.4493702,
final_local_scores=0.4771871, pem=0.0, similarity=18.329927]
3: 6573050 Flamurtari_Vlorë
[pem_pos=0, final_scores=0.69344044,
final_local_scores=0.70993793, pem=0.0, similarity=25.275677]
4: 738 Albania
[pem_pos=0, final_scores=0.5078744,
final_local_scores=0.52569324, pem=-0.41162592,
similarity=22.015934]
5: 26830 Slovakia
[pem_pos=0, final_scores=0.16784279,
final_local_scores=0.13156801, pem=-0.36602598,
similarity=10.006471]
6: 415846 Bistrița
[pem_pos=0, final_scores=0.44542444,
final_local_scores=0.53998834, pem=0.0, similarity=20.204042]
7: 2384775 ACF_Gloria_Bistrița
[pem_pos=0, final_scores=0.6722526,
final_local_scores=0.7054546, pem=0.0, similarity=25.141886]
8: 25445 Romania
[pem_pos=0, final_scores=0.12063189,
final_local_scores=0.10426569, pem=-0.21241479,
similarity=8.356354]
9: 2990151 Valletta_F.C.
[pem_pos=1, final_scores=0.5535626,
final_local_scores=0.59742725, pem=-0.9986768,

```
similarity=27.349092]
10: 1009065 Malta_national_football_team
    [pem_pos=1, final_scores=0.49120933,
final_local_scores=0.49350816, pem=-2.810524,
similarity=34.101048]
11: 2384775 ACF_Gloria_Bistrița
    [pem_pos=0, final_scores=0.9561487,
final_local_scores=1.015568, pem=0.0, similarity=34.396286]
12: 2990151 Valletta_F.C.
    [pem_pos=1, final_scores=0.6416977,
final_local_scores=0.69370145, pem=-0.9986768,
similarity=30.222107]
13: 10424084 Gilbert_Agius
    [pem_pos=0, final_scores=0.53906244,
final_local_scores=0.64191765, pem=0.0, similarity=23.245815]
14: 2384775 ACF_Gloria_Bistrița
    [pem_pos=0, final_scores=0.6699424,
final_local_scores=0.70293105, pem=0.0, similarity=25.066578]
15: 74612 Chorzów
    [pem_pos=0, final_scores=0.4302851,
final_local_scores=0.49114, pem=0.0, similarity=18.746311]
16: 677685 Ruch_Chorzów
    [pem_pos=0, final_scores=0.67252475,
final_local_scores=0.70780295, pem=0.0, similarity=25.211966]
17: 771860 Poland_national_football_team
    [pem_pos=1, final_scores=0.5803104,
final_local_scores=0.5708483, pem=-3.319454,
similarity=39.17667]
18: 300136 The_New_Saints_F.C.
    [pem_pos=1, final_scores=0.4190939,
final_local_scores=0.44890255, pem=-0.4780358,
similarity=20.085493]
19: 580291 Wales_national_football_team
    [pem_pos=1, final_scores=0.27238965,
final_local_scores=0.25554094, pem=-2.4526134,
similarity=25.053257]
20: 2996295 Arkadiusz_Bąk
    [pem_pos=0, final_scores=0.46076232,
final_local_scores=0.58211946, pem=0.0, similarity=21.461319]
21: 677685 Ruch_Chorzów
    [pem_pos=0, final_scores=0.4657541,
final_local_scores=0.4819374, pem=0.0, similarity=18.471687]
22: 287852 Larnaca
    [pem_pos=0, final_scores=0.36361712,
final_local_scores=0.40599877, pem=0.0, similarity=16.205526]
23: 3550878 AEK_Larnaca_F.C.
    [pem_pos=0, final_scores=0.38514692,
final_local_scores=0.39402562, pem=0.0, similarity=15.848223]
24: 722717 Cyprus_national_football_team
    [pem_pos=1, final_scores=0.45960873,
final_local_scores=0.44797683, pem=-2.9483724,
similarity=33.491943]
```

25: 695292 Armenia_national_football_team
[pem_pos=1, final_scores=0.417445,
final_local_scores=0.4045248, pem=-2.6870568,
similarity=30.774174]
26: 3550878 AEK_Larnaca_F.C.
[pem_pos=0, final_scores=0.38662368,
final_local_scores=0.3956387, pem=0.0, similarity=15.89636]
27: 404466 Šiauliai
[pem_pos=0, final_scores=0.3925851,
final_local_scores=0.4544403, pem=0.0, similarity=17.651117]
28: 6492937 FK_Kareda_Kaunas
[pem_pos=0, final_scores=0.40590614,
final_local_scores=0.50091326, pem=0.0, similarity=19.037964]
29: 17675 Lithuania
[pem_pos=0, final_scores=0.29424244,
final_local_scores=0.28596056, pem=-0.39766282,
similarity=14.785898]
30: 1061075 FC_Sion
[pem_pos=1, final_scores=0.71650547,
final_local_scores=0.7410267, pem=-1.4381809,
similarity=34.02448]
31: 681862 Switzerland_national_football_team
[pem_pos=1, final_scores=0.55583113,
final_local_scores=0.55299455, pem=-2.6520712,
similarity=35.01455]
32: 1061075 FC_Sion
[pem_pos=1, final_scores=0.90243644,
final_local_scores=0.944128, pem=-1.4381809,
similarity=40.085426]
33: 1213985 Vinnytsia
[pem_pos=0, final_scores=0.29965138,
final_local_scores=0.35478854, pem=0.0, similarity=14.677311]
34: 1008682 Ukraine_national_football_team
[pem_pos=2, final_scores=0.37680268,
final_local_scores=0.35239774, pem=-3.0947542,
similarity=31.435719]
35: 6680744 JK_Tallinna_Sadam
[pem_pos=0, final_scores=0.40278953,
final_local_scores=0.5319368, pem=0.0, similarity=19.963768]
36: 28222445 Estonia
[pem_pos=0, final_scores=0.43444067,
final_local_scores=0.45981872, pem=-0.4106629,
similarity=20.04487]
37: 7010481 PFC_Nyva_Vinnytsia
[pem_pos=0, final_scores=0.21838717,
final_local_scores=0.25744706, pem=0.0, similarity=11.772447]
38: 56494 Bergen
[pem_pos=0, final_scores=0.57039964,
final_local_scores=0.6420257, pem=-0.28360042,
similarity=24.791304]
39: 192068 SK_Brann
[pem_pos=0, final_scores=0.6052528,

```
final_local_scores=0.6154682, pem=-0.031611566,
similarity=22.62842]
40: 21241 Norway
    [pem_pos=0, final_scores=0.4890678,
final_local_scores=0.5294595, pem=-0.22668563,
similarity=21.122591]
41: 830612 Shelbourne_F.C.
    [pem_pos=0, final_scores=0.51182055,
final_local_scores=0.52560294, pem=0.0, similarity=19.774754]
42: 147575 Ireland
    [pem_pos=0, final_scores=0.11089258,
final_local_scores=0.12946633, pem=-0.5653214,
similarity=11.027552]
43: 192068 SK_Brann
    [pem_pos=0, final_scores=0.73457557,
final_local_scores=0.75673366, pem=-0.031611566,
similarity=26.844063]
44: 3118127 Mons_Ivar_Mjelde
    [pem_pos=0, final_scores=0.5035512,
final_local_scores=0.5773474, pem=0.0, similarity=21.31891]
45: 13926052 Jan_Ove_Pedersen
    [pem_pos=0, final_scores=0.55002016,
final_local_scores=0.6288481, pem=0.0, similarity=22.855793]
46: 830612 Shelbourne_F.C.
    [pem_pos=0, final_scores=0.60690933,
final_local_scores=0.629473, pem=0.0, similarity=22.874441]
47: 6056077 Mark_Rutherford_(footballer)
    [pem_pos=0, final_scores=0.59726,
final_local_scores=0.70233035, pem=-0.26290208,
similarity=26.478354]
48: 192068 SK_Brann
    [pem_pos=0, final_scores=0.6810549,
final_local_scores=0.6982704, pem=-0.031611566,
similarity=25.099401]
49: 57644 Sofia
    [pem_pos=0, final_scores=0.59271127,
final_local_scores=0.61373687, pem=-0.10680699,
similarity=22.985678]
50: 893469 PFC_Levski_Sofia
    [pem_pos=0, final_scores=0.47316116,
final_local_scores=0.45370603, pem=0.0, similarity=17.629206]
51: 3415 Bulgaria
    [pem_pos=0, final_scores=0.434161,
final_local_scores=0.45249426, pem=-0.31727514,
similarity=19.318436]
52: 27792964 NK_Olimpija_Ljubljana_(defunct)
    [pem_pos=1, final_scores=0.26263243,
final_local_scores=0.23484135, pem=-0.87113035,
similarity=15.835189]
53: 27338 Slovenia
    [pem_pos=0, final_scores=0.27626717,
final_local_scores=0.2783656, pem=-0.27845925,
```

```
similarity=13.911003]
54: 27792964 NK_Olimpija_Ljubljana_(defunct)
    [pem_pos=1, final_scores=0.18987103,
final_local_scores=0.15536058, pem=-0.87113035,
similarity=13.4633255]
55: 754336 Rheinpark_Stadion
    [pem_pos=2, final_scores=0.12409465,
final_local_scores=0.14462155, pem=-1.0907357,
similarity=14.337098]
56: 827884 FC_Vaduz
    [pem_pos=1, final_scores=0.46186787,
final_local_scores=0.4696589, pem=-0.75067395,
similarity=22.187553]
57: 652703 Liechtenstein_national_football_team
    [pem_pos=1, final_scores=0.48124927,
final_local_scores=0.48028082, pem=-1.9919682,
similarity=29.254883]
58: 17514 Latvia
    [pem_pos=0, final_scores=0.40527928,
final_local_scores=0.40811288, pem=-0.35144013,
similarity=18.179802]
59: 827884 FC_Vaduz
    [pem_pos=1, final_scores=0.4969253,
final_local_scores=0.50795376, pem=-0.75067395,
similarity=23.330347]
60: 827884 FC_Vaduz
    [pem_pos=1, final_scores=0.43558478,
final_local_scores=0.4409486, pem=-0.75067395,
similarity=21.33078]
61: 57035 Luxembourg_(city)
    [pem_pos=2, final_scores=0.16181217,
final_local_scores=0.18014485, pem=-3.0681052,
similarity=26.150427]
62: 5337201 Union_Luxembourg
    [pem_pos=0, final_scores=0.307666,
final_local_scores=0.37926555, pem=0.0, similarity=15.407753]
63: 17515 Luxembourg
    [pem_pos=0, final_scores=0.3990907,
final_local_scores=0.4076969, pem=-0.49845144,
similarity=18.966858]
64: 715240 Croatia_national_football_team
    [pem_pos=1, final_scores=0.30880278,
final_local_scores=0.266477, pem=-2.4638896,
similarity=25.440933]
65: 43979 Tórshavn
    [pem_pos=0, final_scores=0.3935671,
final_local_scores=0.4705665, pem=0.0, similarity=18.132357]
66: 831312 Havnar_Bóltfelag
    [pem_pos=0, final_scores=0.5679588,
final_local_scores=0.6306505, pem=0.0, similarity=22.909582]
67: 1011425 Faroe_Islands_national_football_team
    [pem_pos=1, final_scores=0.43719715,
```

```
final_local_scores=0.42990714, pem=-0.8728211,  
similarity=21.665535]  
68: 6581309 FC_Dinamo_Batumi  
    [pem_pos=0, final_scores=0.3378486,  
final_local_scores=0.40577877, pem=0.0, similarity=16.198961]  
69: 48768 Georgia_(country)  
    [pem_pos=1, final_scores=0.36414093,  
final_local_scores=0.39859992, pem=-1.3466994,  
similarity=23.30829]  
70: 6581309 FC_Dinamo_Batumi  
    [pem_pos=0, final_scores=0.31505913,  
final_local_scores=0.3808847, pem=0.0, similarity=15.456073]  
71: 23844 Prague  
    [pem_pos=0, final_scores=0.56292444,  
final_local_scores=0.60931903, pem=-0.05392049,  
similarity=22.566236]  
72: 861331 AC_Sparta_Prague  
    [pem_pos=0, final_scores=0.6019334,  
final_local_scores=0.5912471, pem=0.0, similarity=21.733706]  
73: 729739 Czech_Republic_national_football_team  
    [pem_pos=1, final_scores=0.3477047,  
final_local_scores=0.31451964, pem=-2.7919292,  
similarity=28.658552]  
74: 828172 Glentoran_F.C.  
    [pem_pos=0, final_scores=0.4312201,  
final_local_scores=0.4353869, pem=0.0, similarity=17.082527]  
75: 426388 Northern_Ireland_national_football_team  
    [pem_pos=1, final_scores=0.35807484,  
final_local_scores=0.34719446, pem=-2.210495,  
similarity=26.471703]  
76: 5363309 Horst_Siegl  
    [pem_pos=0, final_scores=0.5906315,  
final_local_scores=0.70112586, pem=0.0, similarity=25.012707]  
77: 12633520 Zdeněk_Svoboda  
    [pem_pos=0, final_scores=0.5675677,  
final_local_scores=0.6681199, pem=0.0, similarity=24.027744]  
78: 5364238 Petr_Gabriel  
    [pem_pos=0, final_scores=0.5175979,  
final_local_scores=0.62733436, pem=0.0, similarity=22.810621]  
79: 9602 Edinburgh  
    [pem_pos=0, final_scores=0.5767028,  
final_local_scores=0.649771, pem=-0.11916062,  
similarity=24.128187]  
80: 302212 Heart_of_Midlothian_F.C.  
    [pem_pos=0, final_scores=0.5191765,  
final_local_scores=0.5135193, pem=-0.29860955,  
similarity=21.038038]  
81: 26994 Scotland  
    [pem_pos=0, final_scores=0.2839991,  
final_local_scores=0.32050726, pem=-0.22905664,  
similarity=14.899935]  
82: 361966 Red_Star_Belgrade
```

[pem_pos=0, final_scores=0.4138586,
final_local_scores=0.38506478, pem=0.0, similarity=15.580813]
83: 34244 Yugoslavia
[pem_pos=0, final_scores=0.20398806,
final_local_scores=0.18985033, pem=-0.6536551,
similarity=13.309903]
84: 302212 Heart_of_Midlothian_F.C.
[pem_pos=0, final_scores=0.64203006,
final_local_scores=0.6477182, pem=-0.29860955,
similarity=25.0428]
85: 6789447 Dave_McPherson_(footballer)
[pem_pos=1, final_scores=0.6372101,
final_local_scores=0.73847777, pem=-0.16726951,
similarity=27.037]
86: 361966 Red_Star_Belgrade
[pem_pos=0, final_scores=0.4105699,
final_local_scores=0.38147235, pem=-1.0362908,
similarity=21.10912]
87: 23671907 Vinko_Marinović
[pem_pos=0, final_scores=0.6614263,
final_local_scores=0.73822385, pem=0.0, similarity=26.119785]
88: 361966 Red_Star_Belgrade
[pem_pos=0, final_scores=0.3551126,
final_local_scores=0.32089365, pem=-1.0362908,
similarity=19.30133]
89: 5897396 Hapoel_Rishon_LeZion_F.C.
[pem_pos=0, final_scores=0.3447075,
final_local_scores=0.38523078, pem=0.0, similarity=15.585768]
90: 9282173 Israel
[pem_pos=0, final_scores=0.3452235,
final_local_scores=0.40301007, pem=-0.21279564,
similarity=17.273554]
91: 833887 FC_Tiraspol
[pem_pos=0, final_scores=0.4439841,
final_local_scores=0.46678495, pem=0.0, similarity=18.019506]
92: 1011292 Moldova_national_football_team
[pem_pos=1, final_scores=0.3536448,
final_local_scores=0.34615505, pem=-2.4170835,
similarity=27.564146]
93: 2079306 MYPA
[pem_pos=0, final_scores=0.46958065,
final_local_scores=0.5222811, pem=0.0, similarity=19.675623]
94: 10577 Finland
[pem_pos=0, final_scores=0.2006561,
final_local_scores=0.19272596, pem=-0.2445894,
similarity=11.171155]
95: 655906 Azerbaijan_national_football_team
[pem_pos=1, final_scores=0.32714528,
final_local_scores=0.30783, pem=-2.6132543,
similarity=27.487257]
96: 29618 Skopje
[pem_pos=0, final_scores=0.45239967,

```
final_local_scores=0.46405995, pem=0.0, similarity=17.938187]
97: 23564616 Republic_of_Macedonia
    [pem_pos=0, final_scores=0.3262571,
final_local_scores=0.32255167, pem=-1.1139666,
similarity=19.773222]
98: 833965 Budapest_Honvéd_FC
    [pem_pos=0, final_scores=0.55299336,
final_local_scores=0.54373443, pem=0.0, similarity=20.315832]
99: 13275 Hungary
    [pem_pos=0, final_scores=0.35614818,
final_local_scores=0.35374835, pem=-0.32577527,
similarity=16.417885]
100: 833965 Budapest_Honvéd_FC
    [pem_pos=0, final_scores=0.46807653,
final_local_scores=0.45097572, pem=0.0, similarity=17.547728]
101: 833887 FC_Tiraspol
    [pem_pos=0, final_scores=0.4685856,
final_local_scores=0.49365836, pem=0.0, similarity=18.821463]
```

Sample: b'45_SOCCER'

```
SOCCER - [1 1 GOTHENBURG 1] PUT [2 2 FERENCVAROS 2] OUT OF [3
3 EURO CUP 3] . [4 4 BUDAPEST 4] 1996-08-21 [5 5 IFK Gothenbu
rg 5] of [6 6 Sweden 6] drew 1-1 ( 1-0 ) with [7 7 Ferencvaros
7] of [8 8 Hungary 8] in the second leg of their [9 9 European
Champions Cup 9] preliminary round tie played on Wednesday . [
10 10 Gothenburg 10] go through 4-1 on aggregate . Scorers : [
11 11 Ferencvaros 11] : [12 12 Ferenc Horvath 12] ( 15th ) [13
13 IFK Gothenburg 13] : [14 14 Andreas Andersson 14] ( 87th )
Attendance : 9,000
1: 475370 IFK_Göteborg
2: 861261 Ferencvárosi_TC
3: 249510 UEFA_European_Championship
4: 36787 Budapest
5: 475370 IFK_Göteborg
6: 5058739 Sweden
7: 861261 Ferencvárosi_TC
8: 13275 Hungary
9: 44220 UEFA_Champions_League
10: 475370 IFK_Göteborg
11: 861261 Ferencvárosi_TC
12: 6810641 Ferenc_Horváth
13: 475370 IFK_Göteborg
14: 2918991 Andreas_Andersson
1: 475370 IFK_Göteborg
    [pem_pos=3, final_scores=0.55973476,
final_local_scores=0.5334296, pem=-2.4951115,
similarity=33.57712]
2: 861261 Ferencvárosi_TC
    [pem_pos=0, final_scores=0.697163,
final_local_scores=0.6860058, pem=0.0, similarity=24.561493]
3: 249510 UEFA_European_Championship
    [pem_pos=0, final_scores=0.3844741,
```

```
final_local_scores=0.37628472, pem=0.0, similarity=15.318799]
4: 36787 Budapest
    [pem_pos=0, final_scores=0.5179791,
final_local_scores=0.5628125, pem=-0.030229695,
similarity=21.049553]
5: 475370 IFK_Göteborg
    [pem_pos=0, final_scores=0.6443307,
final_local_scores=0.6258378, pem=0.0, similarity=22.76596]
6: 5058739 Sweden
    [pem_pos=0, final_scores=0.51021,
final_local_scores=0.5386703, pem=-0.2135508,
similarity=21.32603]
7: 861261 Ferencvárosi_TC
    [pem_pos=0, final_scores=0.46980244,
final_local_scores=0.4376489, pem=0.0, similarity=17.150028]
8: 13275 Hungary
    [pem_pos=0, final_scores=0.41920596,
final_local_scores=0.45186687, pem=-0.32577527,
similarity=19.34594]
9: 44220 UEFA_Champions_League
    [pem_pos=0, final_scores=0.4557709,
final_local_scores=0.43596417, pem=-0.2526813,
similarity=18.473875]
10: 475370 IFK_Göteborg
    [pem_pos=3, final_scores=0.56198317,
final_local_scores=0.5358857, pem=-2.4951115,
similarity=33.650414]
11: 861261 Ferencvárosi_TC
    [pem_pos=0, final_scores=0.42598963,
final_local_scores=0.38979006, pem=0.0, similarity=15.721826]
12: 6810641 Ferenc_Horváth
    [pem_pos=0, final_scores=0.47629166,
final_local_scores=0.60418606, pem=0.0, similarity=22.11983]
13: 475370 IFK_Göteborg
    [pem_pos=0, final_scores=0.55861455,
final_local_scores=0.53220594, pem=0.0, similarity=19.9718]
14: 2918991 Andreas_Andersson
    [pem_pos=0, final_scores=0.47913748,
final_local_scores=0.56771183, pem=0.0, similarity=21.031366]
```

Sample: b'46_SOCCER'

```
SOCCER - [1 1 BRAZILIAN 1] CHAMPIONSHIP RESULTS . [2 2 RIO DE JANEIRO 2] 1996-08-22 Results of midweek matches in the [3 3 Brazilian 3] soccer championship . [4 4 Bahia 4] 2 [5 5 Atletico Paranaense 5] 0 [6 6 Corinthians 6] 1 [7 7 Guarani 7] 0 [8 8 Coritiba 8] 1 [9 9 Atletico Mineiro 9] 0 [10 10 Cruzeiro 10] 2 [11 11 Vitoria 11] 1 [1 1 Flamengo 1] 0 [12 12 Juventude 12] 1 [13 13 Goias 13] 3 [14 14 Sport Recife 14] 1 [15 15 Gremio 15] 6 [16 16 Bragantino 16] 1 [17 17 Palmeiras 17] 3 [2 2 Vasco da Gama 2] 1 [18 18 Portuguesa 18] 2 [19 19 Parana 19] 0
```

1: 3383 Brazil

2: 25936 Rio_de_Janeiro

```
3: 3383 Brazil
4: 187328 Esporte_Clube_Bahia
5: 187346 Clube_Atlético_Paranaense
6: 187241 Sport_Club_Corinthians_Paulista
7: 1367196 Guarani_Futebol_Clube
8: 1550341 Coritiba_Foot_Ball_Club
9: 187305 Clube_Atlético_Mineiro
10: 187301 Cruzeiro_Esporte_Clube
11: 187310 Esporte_Clube_Vitória
12: 1577507 Esporte_Clube_Juventude
13: 1577470 Goiás_Esporte_Clube
14: 604575 Sport_Club_do_Recife
15: 18564144 Grêmio_Foot-Ball_Porto_Alegrense
16: 1861824 Clube_Atlético_Bragantino
17: 187243 Sociedade_Esportiva_Palmeiras
18: 1742904 Associação_Portuguesa_de_Desportos
19: 311141 Paraná_Clube
1: 21682157 Esporte_Clube_Flamengo
    [winner: 187239 Clube_de_Regatas_do_Flamengo, pem_pos=0,
final_scores=0.31433076, final_local_scores=0.24273556,
pem=0.0, similarity=11.333426]
    gt: 21682157 Esporte_Clube_Flamengo, pem_pos=2,
final_scores=-0.07361989, final_local_scores=-0.087050706,
pem=-2.2767625, similarity=13.873335 ]
2: 8805153 Associação_Desportiva_Vasco_da_Gama
    [winner: 187324 CR_Vasco_da_Gama, pem_pos=1,
final_scores=0.26205987, final_local_scores=0.18721709,
pem=-0.725215, similarity=13.620476]
    gt: 8805153 Associação_Desportiva_Vasco_da_Gama,
pem_pos=5, final_scores=-0.049567983,
final_local_scores=-0.01806596, pem=-1.927182,
similarity=14.030904 ]
1: 3383 Brazil
    [pem_pos=0, final_scores=0.738343,
final_local_scores=0.8070483, pem=-0.2733083,
similarity=29.659937]
2: 25936 Rio_de_Janeiro
    [pem_pos=0, final_scores=0.4601344,
final_local_scores=0.4837039, pem=0.0, similarity=18.524403]
3: 3383 Brazil
    [pem_pos=0, final_scores=0.7484662,
final_local_scores=0.81810635, pem=-0.2733083,
similarity=29.989931]
4: 187328 Esporte_Clube_Bahia
    [pem_pos=1, final_scores=0.31282812,
final_local_scores=0.24119654, pem=-1.7376727,
similarity=20.737234]
5: 187346 Clube_Atlético_Paranaense
    [pem_pos=0, final_scores=0.3977303,
final_local_scores=0.34143436, pem=0.0, similarity=14.278795]
6: 187241 Sport_Club_Corinthians_Paulista
    [pem_pos=0, final_scores=0.4447456,
```

```
final_local_scores=0.38337737, pem=-0.017981783,
similarity=15.628245]
7: 1367196 Guarani_Futebol_Clube
    [pem_pos=0, final_scores=0.3694117,
final_local_scores=0.31433344, pem=-0.25016624,
similarity=14.830492]
8: 1550341 Coritiba_Foot_Ball_Club
    [pem_pos=0, final_scores=0.5228583,
final_local_scores=0.48261815, pem=0.0, similarity=18.492]
9: 187305 Clube_Atlético_Mineiro
    [pem_pos=0, final_scores=0.35611874,
final_local_scores=0.28969958, pem=0.0, similarity=12.734926]
10: 187301 Cruzeiro_Esporte_Clube
    [pem_pos=0, final_scores=0.44360054,
final_local_scores=0.38667428, pem=0.0, similarity=15.628844]
11: 187310 Esporte_Clube_Vitória
    [pem_pos=2, final_scores=0.3490243,
final_local_scores=0.2855145, pem=-1.938656,
similarity=23.152748]
12: 1577507 Esporte_Clube_Juventude
    [pem_pos=0, final_scores=0.46287632,
final_local_scores=0.42873532, pem=0.0, similarity=16.88403]
13: 1577470 Goiás_Esporte_Clube
    [pem_pos=1, final_scores=0.42811167,
final_local_scores=0.3901112, pem=-0.957059,
similarity=20.936045]
14: 604575 Sport_Club_do_Recife
    [pem_pos=0, final_scores=0.49094844,
final_local_scores=0.44524145, pem=0.0, similarity=17.376606]
15: 18564144 Grêmio_Foot-Ball_Porto_Alegrense
    [pem_pos=0, final_scores=0.5237759,
final_local_scores=0.47306406, pem=0.0, similarity=18.206888]
16: 1861824 Clube_Atlético_Bragantino
    [pem_pos=0, final_scores=0.35659516,
final_local_scores=0.336656, pem=0.0, similarity=14.1362]
17: 187243 Sociedade_Esportiva_Palmeiras
    [pem_pos=0, final_scores=0.44049376,
final_local_scores=0.38045818, pem=0.0, similarity=15.443344]
18: 1742904 Associação_Portuguesa_de_Desportos
    [pem_pos=0, final_scores=0.3359949,
final_local_scores=0.27797684, pem=-0.14295863,
similarity=13.162527]
19: 311141 Paraná_Clube
    [pem_pos=4, final_scores=0.47886568,
final_local_scores=0.45283824, pem=-1.6926686,
similarity=26.808306]
1: 187239 Clube_de_Regatas_do_Flamengo
    [pem_pos=0, final_scores=0.31433076,
final_local_scores=0.24273556, pem=0.0, similarity=11.333426]
2: 187324 CR_Vasco_da_Gama
    [pem_pos=1, final_scores=0.26205987,
final_local_scores=0.18721709, pem=-0.725215,
```

similarity=13.620476]

Sample: b'47_TENNIS'

TENNIS - [1 1 NEWCOMBE 1] PONDERS HIS [2 2 DAVIS CUP 2]
FUTURE . [3 3 SYDNEY 3] 1996-08-22 [4 4 Australian 4] [5 5 Da
vis Cup 5] captain [6 6 John Newcombe 6] on Thursday
signalled his possible resignation if his team loses an away
tie against [7 7 Croatia 7] next month . The former [8 8 Wimb
ledon 8] champion said the immediate future of [9 9 Australia
9] 's [10 10 Davis Cup 10] coach [11 11 Tony Roche 11] could
also be determined by events in [12 12 Split 12] . " If we
lose this one , [13 13 Tony 13] and I will have to have a
good look at giving someone else a go , " [14 14 Newcombe 14]
was quoted as saying in [15 15 Sydney 15] 's [16 16 Daily Tel
egraph 16] newspaper . [17 17 Australia 17] face [18 18 Croat
ia 18] in the world group qualifying tie on clay from
September 20-22 . Under [19 19 Newcombe 19] 's leadership , [2
0 20 Australia 20] were relegated from the elite world group
last year , the first time the 26-time [21 21 Davis Cup 21]
winners had slipped from the top rank . Since taking over as
captain from [22 22 Neale Fraser 22] in 1994 , [23 23 Newcombe
23] 's record in tandem with [24 24 Roche 24] , his former
doubles partner , has been three wins and three losses . [25 2
5 Newcombe 25] has selected [26 26 Wimbledon 26] semifinalist
[27 27 Jason Stoltenberg 27] , [28 28 Patrick Rafter 28] , [29
29 Mark Philippoussis 29] , and [30 30 Olympic 30] doubles
champions [31 31 Todd Woodbridge 31] and [32 32 Mark Woodforde
32] to face the Croatians . The home side boasts world number
six [33 33 Goran Ivanisevic 33] , and [34 34 Newcombe 34]
conceded his players would be hard-pressed to beat the
Croatian number one . " We are ready to fight to our last
breath -- [35 35 Australia 35] must play at its absolute best
to win , " said [36 36 Newcombe 36] , who described the tie
as the toughest he has faced as captain . [37 37 Australia 37]
last won the [38 38 Davis Cup 38] in 1986 , but they were
beaten finalists against [39 39 Germany 39] three years ago
under [40 40 Fraser 40] 's guidance .

1: 481779 John_Newcombe
2: 227983 Davis_Cup
3: 27862 Sydney
4: 4689264 Australia
5: 227983 Davis_Cup
6: 481779 John_Newcombe
7: 9663393 Croatia_Davis_Cup_team
8: 91195 The_Championships,_Wimbledon
9: 4689264 Australia
10: 227983 Davis_Cup
11: 870020 Tony_Roche
12: 18952408 Split,_Croatia
13: 870020 Tony_Roche
14: 481779 John_Newcombe
15: 27862 Sydney

```
16: 172311 The_Daily_Telegraph_(Australia)
17: 9656856 Australia_Davis_Cup_team
18: 9663393 Croatia_Davis_Cup_team
19: 481779 John_Newcombe
20: 9656856 Australia_Davis_Cup_team
21: 227983 Davis_Cup
22: 2199552 Neale_Fraser
23: 481779 John_Newcombe
24: 870020 Tony_Roche
25: 481779 John_Newcombe
26: 91195 The_Championships,_Wimbledon
27: 2598440 Jason_Stoltenberg
28: 23297 Patrick_Rafter
29: 261368 Mark_Philippoussis
30: 27566 Summer_Olympic_Games
31: 1547951 Todd_Woodbridge
32: 1918850 Mark_Woodforde
33: 194244 Goran_Ivanišević
34: 481779 John_Newcombe
35: 9656856 Australia_Davis_Cup_team
36: 481779 John_Newcombe
37: 9656856 Australia_Davis_Cup_team
38: 227983 Davis_Cup
39: 9663079 Germany_Davis_Cup_team
40: 2199552 Neale_Fraser
1: 481779 John_Newcombe
    [pem_pos=0, final_scores=0.73969406,
final_local_scores=0.716238, pem=0.0, similarity=25.463684]
2: 227983 Davis_Cup
    [pem_pos=0, final_scores=0.549179,
final_local_scores=0.5303175, pem=-0.21573581,
similarity=21.08865]
3: 27862 Sydney
    [pem_pos=0, final_scores=0.5293285,
final_local_scores=0.5913588, pem=-0.1443489,
similarity=22.522032]
4: 4689264 Australia
    [pem_pos=0, final_scores=0.51254493,
final_local_scores=0.60439736, pem=-0.491009,
similarity=24.796318]
5: 227983 Davis_Cup
    [pem_pos=0, final_scores=0.4191664,
final_local_scores=0.38829845, pem=-0.21573581,
similarity=16.85052]
6: 481779 John_Newcombe
    [pem_pos=0, final_scores=0.6964314,
final_local_scores=0.6672539, pem=0.0, similarity=24.0019]
7: 9663393 Croatia_Davis_Cup_team
    [pem_pos=3, final_scores=0.33457464,
final_local_scores=0.38533586, pem=-3.7709618,
similarity=36.095978]
8: 91195 The_Championships,_Wimbledon
```

```
[pem_pos=0, final_scores=0.59058636,
final_local_scores=0.56593126, pem=-0.8505555,
similarity=25.603683]
9: 4689264 Australia
    [pem_pos=0, final_scores=0.5858314,
final_local_scores=0.6844517, pem=-0.17168477,
similarity=25.448767]
10: 227983 Davis_Cup
    [pem_pos=0, final_scores=0.45070106,
final_local_scores=0.4227453, pem=-0.21573581,
similarity=17.87848]
11: 870020 Tony_Roche
    [pem_pos=0, final_scores=0.57912415,
final_local_scores=0.5727002, pem=0.0, similarity=21.18023]
12: 18952408 Split,_Croatia
    [pem_pos=0, final_scores=0.38307238,
final_local_scores=0.48449796, pem=-0.6543379,
similarity=22.10649]
13: 870020 Tony_Roche
    [pem_pos=0, final_scores=1.0053322,
final_local_scores=1.0382679, pem=0.0, similarity=35.0737]
14: 481779 John_Newcombe
    [pem_pos=0, final_scores=0.80329674,
final_local_scores=0.78571427, pem=0.0, similarity=27.536993]
15: 27862 Sydney
    [pem_pos=0, final_scores=0.4336164,
final_local_scores=0.48680788, pem=-0.1443489,
similarity=19.402023]
16: 172311 The_Daily_Telegraph_(Australia)
    [pem_pos=1, final_scores=0.21845566,
final_local_scores=0.31886634, pem=-0.57183003,
similarity=16.715021]
17: 9656856 Australia_Davis_Cup_team
    [pem_pos=9, final_scores=0.36779797,
final_local_scores=0.36280334, pem=-5.04368,
similarity=42.344807]
18: 9663393 Croatia_Davis_Cup_team
    [pem_pos=3, final_scores=0.32901347,
final_local_scores=0.37926108, pem=-3.7709618,
similarity=35.914696]
19: 481779 John_Newcombe
    [pem_pos=0, final_scores=0.79475886,
final_local_scores=0.77638793, pem=0.0, similarity=27.258677]
20: 9656856 Australia_Davis_Cup_team
    [pem_pos=9, final_scores=0.4641651,
final_local_scores=0.4680698, pem=-5.04368,
similarity=45.486168]
21: 227983 Davis_Cup
    [pem_pos=0, final_scores=0.42541826,
final_local_scores=0.39512765, pem=-0.21573581,
similarity=17.054316]
22: 2199552 Neale_Fraser
```

```
[pem_pos=0, final_scores=0.7953004,
final_local_scores=0.82950634, pem=0.0, similarity=28.843836]
23: 481779 John_Newcombe
    [pem_pos=0, final_scores=0.82556885,
final_local_scores=0.8100432, pem=0.0, similarity=28.263018]
24: 870020 Tony_Roche
    [pem_pos=0, final_scores=1.0863534,
final_local_scores=1.1300066, pem=0.0, similarity=37.811363]
25: 481779 John_Newcombe
    [pem_pos=0, final_scores=0.76446193,
final_local_scores=0.74329317, pem=0.0, similarity=26.271063]
26: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.5712823,
final_local_scores=0.5448445, pem=-0.8505555,
similarity=24.974413]
27: 2598440 Jason_Stoltenberg
    [pem_pos=0, final_scores=0.79578453,
final_local_scores=0.84126467, pem=0.0, similarity=29.194729]
28: 23297 Patrick_Rafter
    [pem_pos=0, final_scores=0.91562325,
final_local_scores=0.95083207, pem=0.0, similarity=32.46444]
29: 261368 Mark_Philippoussis
    [pem_pos=0, final_scores=0.82512754,
final_local_scores=0.8527419, pem=0.0, similarity=29.537231]
30: 27566 Summer_Olympic_Games
    [pem_pos=2, final_scores=0.34776872,
final_local_scores=0.37682384, pem=-3.2532303,
similarity=33.026463]
31: 1547951 Todd_Woodbridge
    [pem_pos=0, final_scores=0.6993922,
final_local_scores=0.7311379, pem=0.0, similarity=25.908325]
32: 1918850 Mark_Woodforde
    [pem_pos=0, final_scores=0.6297074,
final_local_scores=0.67335457, pem=0.0, similarity=24.183956]
33: 194244 Goran_Ivanišević
    [pem_pos=0, final_scores=0.94328636,
final_local_scores=0.9962556, pem=0.0, similarity=33.819965]
34: 481779 John_Newcombe
    [pem_pos=0, final_scores=0.6950141,
final_local_scores=0.66743195, pem=0.0, similarity=24.007214]
35: 9656856 Australia_Davis_Cup_team
    [pem_pos=9, final_scores=0.4664514,
final_local_scores=0.47056723, pem=-5.04368,
similarity=45.560696]
36: 481779 John_Newcombe
    [pem_pos=0, final_scores=0.9026113,
final_local_scores=0.8942004, pem=0.0, similarity=30.774437]
37: 9656856 Australia_Davis_Cup_team
    [pem_pos=9, final_scores=0.50754935,
final_local_scores=0.5154605, pem=-5.04368,
similarity=46.9004]
38: 227983 Davis_Cup
```

```

    [pem_pos=0, final_scores=0.42248577,
final_local_scores=0.39192432, pem=-0.21573581,
similarity=16.958721]
39: 9663079 Germany_Davis_Cup_team
    [pem_pos=12, final_scores=0.38498974,
final_local_scores=0.42018127, pem=-5.051402,
similarity=44.09907]
40: 2199552 Neale_Fraser
    [pem_pos=0, final_scores=1.2562795,
final_local_scores=1.3371394, pem=0.0, similarity=43.99262]

```

Sample: b'48_BADMINTON'

BADMINTON - [1 1 MALAYSIAN OPEN 1] RESULTS . [2 2 KUALA LUMPUR 2] 1996-08-22 Results in the [3 3 Malaysian Open 3]
badminton tournament on Thursday (prefix number denotes seeding) : Men 's singles , third round 9/16 - Luo Yigang ([4 4 China 4]) beat Hwang Sun-ho ([5 5 South Korea 5]) 15-3
15-7 Jason Wong ([6 6 Malaysia 6]) beat Abdul Samad Ismail ([7 7 Malaysia 7]) 16-18 15-2 17-14 P. Kantharoopan ([8 8 Malaysia 8]) beat 3/4 - Jeroen Van Dijk ([9 9 Netherlands 9]) 15-11 18-14 Wijaya Indra ([10 10 Indonesia 10]) beat 5/8
- Pang Chen ([11 11 Malaysia 11]) 15-6 6-15 15-7 3/4 - Hu Zhilan ([12 12 China 12]) beat Nunung Subandoro ([13 13 Indonesia 13]) 5-15 18-15 15-6 9/16 - [14 14 Hermawan Susanto 14] ([15 15 Indonesia 15]) beat 1 - [16 16 Fung Permadi 16] ([17 17 Taiwan 17]) 15-8 15-12 Women 's singles 2nd round 1
- [18 18 Wang Chen 18] ([19 19 China 19]) beat Cindana ([20 20 Indonesia 20]) 11-3 Iama ([21 21 Japan 21]) beat Margit Borg ([22 22 Sweden 22]) 11-6 11-6 Sun Jian ([23 23 China 23]) beat Marina Andrievskaqya ([24 24 Sweden 24]) 11-8
11-2 5/8 - [25 25 Meluawati 25] ([26 26 Indonesia 26]) beat Chan Chia Fong ([27 27 Malaysia 27]) 11-6 11-1 [28 28 Gong Zichao 28] ([29 29 China 29]) beat Liu Lufung ([30 30 China 30]) 6-11 11-7 11-3 Zeng Yaqiong ([31 31 China 31]) beat Li Feng ([32 32 New Zealand 32]) 11-9 11-6 5/8 - [33 33 Christine Magnusson 33] ([34 34 Sweden 34]) beat Ishwari Boopathy ([35 35 Malaysia 35]) 11-1 10-12 11-4 2 - [36 36 Zhang Ning 36] ([37 37 China 37]) beat Olivia ([38 38 Indonesia 38]) 11-8 11-6

1: 10145751 Malaysia_Open_(badminton)
2: 16854 Kuala_Lumpur
3: 10145751 Malaysia_Open_(badminton)
4: 5405 China
5: 27019 South_Korea
6: 3607937 Malaysia
7: 3607937 Malaysia
8: 3607937 Malaysia
9: 21148 Netherlands
10: 14579 Indonesia
11: 3607937 Malaysia
12: 5405 China
13: 14579 Indonesia

```
14: 9521761 Hermawan_Susanto
15: 14579 Indonesia
16: 14504169 Fung_Permatadi
17: 25734 Taiwan
18: 927529 Wang_Chen_(badminton)
19: 5405 China
20: 14579 Indonesia
21: 15573 Japan
22: 5058739 Sweden
23: 5405 China
24: 5058739 Sweden
25: 19927785 Meiluawati
26: 14579 Indonesia
27: 3607937 Malaysia
28: 2491066 Gong_Zhichao
29: 5405 China
30: 5405 China
31: 5405 China
32: 4913064 New_Zealand
33: 14938322 Christine_Magnusson
34: 5058739 Sweden
35: 3607937 Malaysia
36: 927560 Zhang_Ning
37: 5405 China
38: 14579 Indonesia
1: 10145751 Malaysia_Open_(badminton)
    [pem_pos=2, final_scores=0.37028188,
final_local_scores=0.44384187, pem=-1.3940765,
similarity=24.916044]
2: 16854 Kuala_Lumpur
    [pem_pos=0, final_scores=0.45379257,
final_local_scores=0.438502, pem=0.0, similarity=17.175488]
3: 10145751 Malaysia_Open_(badminton)
    [pem_pos=2, final_scores=0.36788398,
final_local_scores=0.44122255, pem=-1.3940765,
similarity=24.83788]
4: 5405 China
    [pem_pos=0, final_scores=0.52990156,
final_local_scores=0.49009693, pem=-0.15669827,
similarity=19.567333]
5: 27019 South_Korea
    [pem_pos=0, final_scores=0.57183385,
final_local_scores=0.564222, pem=-0.22449642,
similarity=22.14807]
6: 3607937 Malaysia
    [pem_pos=0, final_scores=0.66735405,
final_local_scores=0.63377994, pem=-0.20223404,
similarity=24.10275]
7: 3607937 Malaysia
    [pem_pos=0, final_scores=0.6591924,
final_local_scores=0.6248646, pem=-0.20223404,
similarity=23.836699]
```

```
8: 3607937 Malaysia
    [pem_pos=0, final_scores=0.69943124,
final_local_scores=0.66881937, pem=-0.20223404,
similarity=25.148397]
9: 21148 Netherlands
    [pem_pos=0, final_scores=0.5989173,
final_local_scores=0.61521643, pem=-0.24940795,
similarity=23.805319]
10: 14579 Indonesia
    [pem_pos=0, final_scores=0.6264927,
final_local_scores=0.5909154, pem=-0.23151328,
similarity=22.982813]
11: 3607937 Malaysia
    [pem_pos=0, final_scores=0.6514389,
final_local_scores=0.61639506, pem=-0.20223404,
similarity=23.58395]
12: 5405 China
    [pem_pos=0, final_scores=0.54518867,
final_local_scores=0.50679576, pem=-0.15669827,
similarity=20.065659]
13: 14579 Indonesia
    [pem_pos=0, final_scores=0.6720014,
final_local_scores=0.64062685, pem=-0.23151328,
similarity=24.466301]
14: 9521761 Hermawan_Susanto
    [pem_pos=0, final_scores=0.7886167,
final_local_scores=0.9136097, pem=0.0, similarity=31.353647]
15: 14579 Indonesia
    [pem_pos=0, final_scores=0.6378841,
final_local_scores=0.60335886, pem=-0.23151328,
similarity=23.35415]
16: 14504169 Fung_Permadi
    [pem_pos=0, final_scores=0.49969488,
final_local_scores=0.6314735, pem=0.0, similarity=22.934141]
17: 25734 Taiwan
    [pem_pos=0, final_scores=0.48690397,
final_local_scores=0.45700228, pem=-0.077886656,
similarity=18.151133]
18: 927529 Wang_Chen_(badminton)
    [pem_pos=0, final_scores=0.53631157,
final_local_scores=0.6567058, pem=-0.19482312,
similarity=24.746601]
19: 5405 China
    [pem_pos=0, final_scores=0.46958548,
final_local_scores=0.42421073, pem=-0.15669827,
similarity=17.601156]
20: 14579 Indonesia
    [pem_pos=0, final_scores=0.69941443,
final_local_scores=0.6705713, pem=-0.23151328,
similarity=25.359905]
21: 15573 Japan
    [pem_pos=0, final_scores=0.66791254,
```

```
final_local_scores=0.66399336, pem=-0.13549899,  
similarity=24.641462]  
22: 5058739 Sweden  
    [pem_pos=0, final_scores=0.6926606,  
final_local_scores=0.7270492, pem=-0.2135508,  
similarity=26.947632]  
23: 5405 China  
    [pem_pos=0, final_scores=0.50345725,  
final_local_scores=0.46121055, pem=-0.15669827,  
similarity=18.705305]  
24: 5058739 Sweden  
    [pem_pos=0, final_scores=0.77173406,  
final_local_scores=0.81342494, pem=-0.2135508,  
similarity=29.525257]  
25: 19927785 Meiluawati  
    [pem_pos=0, final_scores=0.5369094,  
final_local_scores=0.67228985, pem=0.0, similarity=24.152184]  
26: 14579 Indonesia  
    [pem_pos=0, final_scores=0.6625479,  
final_local_scores=0.6303003, pem=-0.23151328,  
similarity=24.158136]  
27: 3607937 Malaysia  
    [pem_pos=0, final_scores=0.63559395,  
final_local_scores=0.59908676, pem=-0.20223404,  
similarity=23.067436]  
28: 2491066 Gong_Zhichao  
    [pem_pos=0, final_scores=0.7716636,  
final_local_scores=0.91556686, pem=0.0, similarity=31.412056]  
29: 5405 China  
    [pem_pos=0, final_scores=0.5233572,  
final_local_scores=0.48294824, pem=-0.15669827,  
similarity=19.354002]  
30: 5405 China  
    [pem_pos=0, final_scores=0.5109485,  
final_local_scores=0.46939355, pem=-0.15669827,  
similarity=18.949503]  
31: 5405 China  
    [pem_pos=0, final_scores=0.5045164,  
final_local_scores=0.46236748, pem=-0.15669827,  
similarity=18.73983]  
32: 4913064 New_Zealand  
    [pem_pos=0, final_scores=0.6040332,  
final_local_scores=0.6214416, pem=-0.23795462,  
similarity=23.928804]  
33: 14938322 Christine_Magnusson  
    [pem_pos=0, final_scores=0.5397858,  
final_local_scores=0.73123205, pem=0.0, similarity=25.911137]  
34: 5058739 Sweden  
    [pem_pos=0, final_scores=0.74135613,  
final_local_scores=0.7802416, pem=-0.2135508,  
similarity=28.535]  
35: 3607937 Malaysia
```

```
[pem_pos=0, final_scores=0.66940457,
final_local_scores=0.6360198, pem=-0.20223404,
similarity=24.169594]
36: 927560 Zhang_Ning
    [pem_pos=0, final_scores=0.55640525,
final_local_scores=0.6359552, pem=0.0, similarity=23.067884]
37: 5405 China
    [pem_pos=0, final_scores=0.49162555,
final_local_scores=0.44828618, pem=-0.15669827,
similarity=18.319616]
38: 14579 Indonesia
    [pem_pos=0, final_scores=0.6543266,
final_local_scores=0.6213198, pem=-0.23151328,
similarity=23.89014]
```

Sample: b'49_TENNIS'

TENNIS - REVISED MEN 'S DRAW FOR [1 1 U.S. OPEN 1] . [2 2 NEW YORK 2] 1996-08-22 Revised singles draw for the [3 3 U.S. Open 3] tennis championships beginning Monday at the [4 4 U.S 4] . National Tennis Centre (prefix denotes seeding) : Men 's Draw 1 - [5 5 Pete Sampras 5] ([6 6 U.S. 6]) vs. [7 7 Adrian Voinea 7] ([8 8 Romania 8]) [9 9 Jiri Novak 9] ([10 10 Czech Republic 10]) vs. qualifier [11 11 Magnus Larsson 11] ([12 12 Sweden 12]) vs. [13 13 Alexander Volkov 13] ([14 14 Russia 14]) [15 15 Mikael Tillstrom 15] ([16 16 Sweden 16]) vs qualifier Qualifier vs. [17 17 Andrei Olhovskiy 17] ([18 18 Russia 18]) [19 19 Mark Woodforde 19] ([20 20 Australia 20]) vs. [21 21 Mark Philippoussis 21] ([22 22 Australia 22]) [23 23 Roberto Carretero 23] ([24 24 Spain 24]) vs. [25 25 Jordi Burillo 25] ([26 26 Spain 26]) [27 27 Francisco Clavet 27] ([28 28 Spain 28]) vs. 16 - [29 29 Cedric Pioline 29] ([30 30 France 30]) ----- 9 - [31 31 Wayne Ferreira 31] ([32 32 South Africa 32]) vs. qualifier [33 33 Karol Kucera 33] ([34 34 Slovakia 34]) vs. [35 35 Jonas Bjorkman 35] ([36 36 Sweden 36]) Qualifier vs. Christian Rudd ([37 37 Norway 37]) [38 38 Alex Corretja 38] ([39 39 Spain 39]) vs. [40 40 Byron Black 40] ([41 41 Zimbabwe 41]) [42 42 David Rikl 42] ([43 43 Czech Republic 43]) vs. [44 44 Hicham Arazi 44] ([45 45 Morocco 45]) [46 46 Sjeng Schalken 46] ([47 47 Netherlands 47]) vs. [48 48 Gilbert Schaller 48] ([49 49 Austria 49]) [50 50 Grant Stafford 50] ([51 51 South Africa 51]) vs. [52 52 Guy Forget 52] ([53 53 France 53]) [54 54 Fernando Meligeni 54] ([55 55 Brazil 55]) vs. 7 - [56 56 Yevgeny Kafelnikov 56] ([57 57 Russia 57]) ----- 4 - [58 58 Goran Ivanisevic 58] ([59 59 Croatia 59]) vs. [60 60 Andrei Chesnokov 60] ([61 61 Russia 61]) [62 62 Scott Draper 62] ([63 63 Australia 63]) vs. [64 64 Galo Blanco 64] ([65 65 Spain 65]) [66 66 Renzo Furlan 66] ([67 67 Italy 67]) vs. [68 68 Thomas Johansson 68] ([69 69 Sweden 69]) Hendrik Dreekman ([70 70 Germany 70]) vs. [71 71 Greg Rusedski 71] ([72 72 Britain 72]) [73 73 Andrei Medvedev 73] ([74 74 Ukraine 74]) vs. [75 75 Jean-

Philippe Fleurian 75] ([76 76 France 76]) [77 77 Jan Kroslak 77] ([78 78 Slovakia 78]) vs. [79 79 Chris Woodruff 79] ([80 80 U.S. 80]) Qualifier vs. [81 81 Petr Korda 81] ([82 82 Czech Republic 82]) [83 83 Bohdan Ulihrach 83] ([84 84 Czech Republic 84]) vs. 14 - Alberto Costa ([85 85 Spain 85]) ----- 12 - [86 86 Todd Martin 86] ([87 87 U.S. 87]) vs. Younnes El Aynaoui ([88 88 Morocco 88]) [89 89 Andrea Gaudenzi 89] ([90 90 Italy 90]) vs. [91 91 Shuzo Matsuoka 91] ([92 92 Japan 92]) [93 93 Doug Flach 93] ([94 94 U.S. 94]) vs. qualifier [95 95 Mats Wilander 95] ([96 96 Sweden 96]) vs. [97 97 Tim Henman 97] ([98 98 Britain 98]) [99 99 Paul Haarhuis 99] ([100 100 Netherlands 100]) vs. [101 101 Michael Joyce 101] ([102 102 U.S. 102]) [103 103 Michael Tebbutt 103] ([104 104 Australia 104]) vs. [105 105 Richey Reneberg 105] ([106 106 U.S. 106]) [107 107 Jonathan Stark 107] ([108 108 U.S. 108]) vs. [109 109 Bernd Karbacher 109] ([110 110 Germany 110]) [111 111 Stefan Edberg 111] ([112 112 Sweden 112]) vs. 5 - [113 113 Richard Krajicek 113] ([114 114 Netherlands 114]) ----- 6 - [115 115 Andre Agassi 115] ([116 116 U.S. 116]) vs. [117 117 Mauricio H adad 117] ([118 118 Colombia 118]) [119 119 Marcos Ondruska 119] ([120 120 South Africa 120]) vs. [121 121 Felix Mantilla 121] ([122 122 Spain 122]) [123 123 Carlos Moya 123] ([124 124 Spain 124]) vs. [125 125 Scott Humphries 125] ([126 126 U.S. 126]) [127 127 Jan Siemerink 127] ([128 128 Netherlands 128]) vs. [129 129 Carl-Uwe Steeb 129] ([130 130 Germany 130]) Qualifier vs. qualifier [131 131 David Wheaton 131] ([132 132 U.S. 132]) vs. [133 133 Kevin Kim 133] ([134 134 U.S. 134]) [135 135 Nicolas Lapentti 135] ([136 136 Ecuador 136]) vs. Alex O'Brien ([137 137 U.S. 137]) [138 138 Karim Alami 138] ([139 139 Morocco 139]) vs. 11 - [140 140 MaliVai Washington 140] ([141 141 U.S. 141]) ----- 13 - [142 142 Thomas Enqvist 142] ([143 143 Sweden 143]) vs. [144 144 Stephane Simian 144] ([145 145 France 145]) [146 146 Guillaume Raoux 146] ([147 147 France 147]) vs. [148 148 Filip Dewulf 148] ([149 149 Belgium 149]) [150 150 Mark Knowles 150] ([151 151 Bahamas 151]) vs. [152 152 Marcelo Filippini 152] ([153 153 Uruguay 153]) [154 154 Todd Woodbridge 154] ([155 155 Australia 155]) vs. qualifier Kris Goossens ([156 156 Belgium 156]) vs. [157 157 Sergi Bruguera 157] ([158 158 Spain 158]) Qualifier vs. [159 159 Michael Stich 159] ([160 160 Germany 160]) Qualifier vs. [161 161 Chuck Adams 161] ([162 162 U.S. 162]) [163 163 Javier Frana 163] ([164 164 Argentina 164]) vs. 3 - [165 165 Thomas Muster 165] ([166 166 Austria 166]) ----- 8 - [167 167 Jim Courier 167] ([168 168 U.S. 168]) vs. [169 169 Javier Sanchez 169] ([170 170 Spain 170]) [171 171 Jim Grabb 171] ([172 172 U.S. 172]) vs. [173 173 Sandon Stolle 173] ([174 174 Australia 174]) [175 175 Patrick Rafter 175] ([176 176 Australia 176]) vs. [177 177 Kenneth Carlsen 177] ([178 178 Denmark 178]) [179 179 Jason Stoltenberg 179] ([180 180 Australia 180]) vs. [181 181 S

tefano Pescosolido 181] ([182 182 Italy 182]) [183 183 Arnaud Boetsch 183] ([184 184 France 184]) vs. [185 185 Nicolas Pereira 185] ([186 186 Venezuela 186]) [187 187 Carlos Costa 187] ([188 188 Spain 188]) vs. [189 189 Magnus Gustafsson 189] ([190 190 Sweden 190]) [191 191 Jeff Tarango 191] ([192 192 U.S. 192]) vs. [193 193 Alex Radulescu 193] ([194 194 Germany 194]) Qualifier vs. 10 - [195 195 Marcelo Rios 195] ([196 196 Chile 196]) ----- 15 - [197 197 Marc Rosset 197] ([198 198 Switzerland 198] vs. [199 199 Jared Palmer 199] ([200 200 U.S. 200]) [201 201 Martin Damm 201] ([202 202 Czech Republic 202]) vs. [203 203 Hernan Gumy 203] ([204 204 Argentina 204]) [205 205 Nicklas Kulti 205] ([206 206 Sweden 206]) vs. [207 207 Jakob Hlasek 207] ([208 208 Switzerland 208]) [209 209 Cecil Mamait 209] ([210 210 U.S. 210]) vs. [211 211 Alberto Berasategui 211] ([212 212 Spain 212]) [213 213 Vince Spadea 213] ([214 214 U.S. 214]) vs. [215 215 Daniel Vacek 215] ([216 216 Czech Republic 216]) [217 217 David Prinosil 217] ([218 218 Germany 218]) vs. qualifier Qualifier vs. [219 219 Tomas Carbonell 219] ([220 220 Spain 220]) Qualifier vs. 2 - [221 221 Michael Chang 221] ([222 222 U.S. 222])

1: 82729 US_Open_(tennis)
2: 645042 New_York_City
3: 82729 US_Open_(tennis)
4: 3434750 United_States
5: 65961 Pete_Sampras
6: 3434750 United_States
7: 4036370 Adrian_Voinea
8: 25445 Romania
9: 1554049 Jiří_Novák
10: 5321 Czech_Republic
11: 1728546 Magnus_Larsson
12: 5058739 Sweden
13: 1856445 Alexander_Vladimirovich_Volkov
14: 25391 Russia
15: 3943969 Mikael_Tillström
16: 5058739 Sweden
17: 3930485 Andrei_Olhovskiy
18: 25391 Russia
19: 1918850 Mark_Woodforde
20: 4689264 Australia
21: 261368 Mark_Philippoussis
22: 4689264 Australia
23: 1723684 Roberto_Carretero
24: 26667 Spain
25: 14891740 Jordi_Burillo
26: 26667 Spain
27: 4076952 Francisco_Clavet
28: 26667 Spain
29: 979076 Cédric_Pioline
30: 5843419 France
31: 1723516 Wayne_Ferreira

32: 17416221 South_Africa
33: 3039956 Karol_Kučera
34: 26830 Slovakia
35: 1964402 Jonas_Björkman
36: 5058739 Sweden
37: 21241 Norway
38: 17733765 Àlex_Corretja
39: 26667 Spain
40: 1869714 Byron_Black
41: 34399 Zimbabwe
42: 10250905 David_Rikl
43: 5321 Czech_Republic
44: 2508144 Hicham_Arazi
45: 19291 Morocco
46: 2493300 Sjeng_Schalken
47: 21148 Netherlands
48: 20856583 Gilbert_Schaller
49: 26964606 Austria
50: 4026564 Grant_Stafford
51: 17416221 South_Africa
52: 1723582 Guy_Forget
53: 5843419 France
54: 2989127 Fernando_Meligeni
55: 3383 Brazil
56: 489963 Yevgeny_Kafelnikov
57: 25391 Russia
58: 194244 Goran_Ivanišević
59: 5573 Croatia
60: 1723532 Andrei_Chесnokов
61: 25391 Russia
62: 3983307 Scott_Draper
63: 4689264 Australia
64: 441573 Galo_Blanco
65: 26667 Spain
66: 3929800 Renzo_Furlan
67: 14532 Italy
68: 1408278 Thomas_Johansson
69: 5058739 Sweden
70: 11867 Germany
71: 517430 Greg_Rusedski
72: 31717 United_Kingdom
73: 702849 Andriy_Medvedev
74: 31750 Ukraine
75: 4026812 Jean-Philippe_Fleurian
76: 5843419 France
77: 3930155 Ján_Krošlák
78: 26830 Slovakia
79: 2515285 Chris_Woodruff
80: 3434750 United_States
81: 1439769 Petr_Korda
82: 5321 Czech_Republic
83: 2805841 Bohdan_Ulihrach

84: 5321 Czech_Republic
85: 26667 Spain
86: 871725 Todd_Martin
87: 3434750 United_States
88: 19291 Morocco
89: 3929637 Andrea_Gaudenzi
90: 14532 Italy
91: 2575718 Shuzo_Matsuoka
92: 15573 Japan
93: 4025956 Doug_Flach
94: 3434750 United_States
95: 200446 Mats_Wilander
96: 5058739 Sweden
97: 262030 Tim_Henman
98: 31717 United_Kingdom
99: 2186352 Paul_Haarhuis
100: 21148 Netherlands
101: 4026170 Michael_Joyce_(tennis)
102: 3434750 United_States
103: 13355036 Michael_Tebbutt
104: 4689264 Australia
105: 871728 Richey_Reneberg
106: 3434750 United_States
107: 3079144 Jonathan_Stark_(tennis)
108: 3434750 United_States
109: 15290626 Bernd_Karbacher
110: 11867 Germany
111: 391932 Stefan_Edberg
112: 5058739 Sweden
113: 161972 Richard_Krajicek
114: 21148 Netherlands
115: 595 Andre_Agassi
116: 3434750 United_States
117: 20694823 Mauricio_Hadad
118: 5222 Colombia
119: 3930413 Marcos_Ondruska
120: 17416221 South_Africa
121: 1723559 Félix_Mantilla_Botella
122: 26667 Spain
123: 669421 Carlos_Moyá
124: 26667 Spain
125: 20867776 Scott_Humphries
126: 3434750 United_States
127: 2876758 Jan_Siemerink
128: 21148 Netherlands
129: 3254714 Carl-Uwe_Staab
130: 11867 Germany
131: 1829027 David_Wheaton
132: 3434750 United_States
133: 2075801 Kevin_Kim
134: 3434750 United_States
135: 2817661 Nicolás_Lapentti

136: 9334 Ecuador
137: 3434750 United_States
138: 3930942 Karim_Alami
139: 19291 Morocco
140: 873616 MaliVai_Washington
141: 3434750 United_States
142: 933709 Thomas_Enqvist
143: 5058739 Sweden
144: 22577800 Stéphane_Simian
145: 5843419 France
146: 3930324 Guillaume_Raoux
147: 5843419 France
148: 3004535 Filip_Dewulf
149: 3343 Belgium
150: 2261491 Mark_Knowles
151: 3451 The_Bahamas
152: 2936118 Marcelo_Filippini
153: 31843 Uruguay
154: 1547951 Todd_Woodbridge
155: 4689264 Australia
156: 3343 Belgium
157: 913625 Sergi_Bruguera
158: 26667 Spain
159: 979088 Michael_Stich
160: 11867 Germany
161: 20930737 Chuck_Adams
162: 3434750 United_States
163: 3930198 Javier_Frana
164: 18951905 Argentina
165: 166170 Thomas_Muster
166: 26964606 Austria
167: 685780 Jim_Courier
168: 3434750 United_States
169: 420527 Javier_Sánchez
170: 26667 Spain
171: 1915197 Jim_Grabb
172: 3434750 United_States
173: 3823337 Sandon_Stolle
174: 4689264 Australia
175: 23297 Patrick_Rafter
176: 4689264 Australia
177: 1804967 Kenneth_Carlsen
178: 76972 Denmark
179: 2598440 Jason_Stoltenberg
180: 4689264 Australia
181: 3930788 Stefano_Pescosolido
182: 14532 Italy
183: 3930278 Arnaud_Boetsch
184: 5843419 France
185: 3929855 Nicolás_Pereira
186: 32374 Venezuela
187: 2785107 Carlos_Costa_(tennis)

188: 26667 Spain
189: 2793059 Magnus_Gustafsson
190: 5058739 Sweden
191: 1862751 Jeff_Tarango
192: 3434750 United_States
193: 4026041 Alex_Rădulescu
194: 11867 Germany
195: 483554 Marcelo_Ríos
196: 5489 Chile
197: 1717291 Marc_Rosset
198: 26748 Switzerland
199: 3056059 Jared_Palmer
200: 3434750 United_States
201: 4026614 Martin_Damm
202: 5321 Czech_Republic
203: 3929963 Hernán_Gumy
204: 18951905 Argentina
205: 2987792 Nicklas_Kulti
206: 5058739 Sweden
207: 1727962 Jakob_Hlasek
208: 26748 Switzerland
209: 10980025 Cecil_Mamiit
210: 3434750 United_States
211: 1728009 Alberto_Berasategui
212: 26667 Spain
213: 1574767 Vince_Spadea
214: 3434750 United_States
215: 3930665 Daniel_Vacek
216: 5321 Czech_Republic
217: 3930698 David_Prinosil
218: 11867 Germany
219: 13833030 Tomás_Carbonell
220: 26667 Spain
221: 391834 Michael_Chang
222: 3434750 United_States
1: 82729 US_Open_(tennis)
 [pem_pos=1, final_scores=0.5071742,
final_local_scores=0.52210784, pem=-1.2370842,
similarity=26.397907]
2: 645042 New_York_City
 [pem_pos=1, final_scores=0.423819,
final_local_scores=0.45947284, pem=-1.7906687,
similarity=27.539234]
3: 82729 US_Open_(tennis)
 [pem_pos=1, final_scores=0.52364635,
final_local_scores=0.5401012, pem=-1.2370842,
similarity=26.934864]
4: 3434750 United_States
 [pem_pos=0, final_scores=0.45636863,
final_local_scores=0.45600492, pem=0.0, similarity=17.69781]
5: 65961 Pete_Sampras
 [pem_pos=0, final_scores=0.7184549,

```
final_local_scores=0.75323164, pem=0.0, similarity=26.567648]
6: 3434750 United_States
    [pem_pos=0, final_scores=0.49522656,
final_local_scores=0.4984513, pem=-0.14075686,
similarity=19.729952]
7: 4036370 Adrian_Voinea
    [pem_pos=0, final_scores=0.66139483,
final_local_scores=0.7828107, pem=0.0, similarity=27.450346]
8: 25445 Romania
    [pem_pos=0, final_scores=0.7718585,
final_local_scores=0.8060595, pem=-0.21241479,
similarity=29.29928]
9: 1554049 Jiří_Novák
    [pem_pos=0, final_scores=0.83560294,
final_local_scores=0.9252179, pem=0.0, similarity=31.700062]
10: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.5240844,
final_local_scores=0.5244104, pem=-0.27794555,
similarity=21.250677]
11: 1728546 Magnus_Larsson
    [pem_pos=0, final_scores=0.6383018,
final_local_scores=0.7144534, pem=0.0, similarity=25.410427]
12: 5058739 Sweden
    [pem_pos=0, final_scores=0.7641044,
final_local_scores=0.7903747, pem=-0.2135508,
similarity=28.837393]
13: 1856445 Alexander_Vladimirovich_Volkov
    [pem_pos=0, final_scores=0.44191468,
final_local_scores=0.5215274, pem=-0.6501985,
similarity=23.189013]
14: 25391 Russia
    [pem_pos=0, final_scores=0.8458068,
final_local_scores=0.90997654, pem=-0.20299491,
similarity=32.349144]
15: 3943969 Mikael_Tillström
    [pem_pos=0, final_scores=0.6943276,
final_local_scores=0.7924459, pem=0.0, similarity=27.737879]
16: 5058739 Sweden
    [pem_pos=0, final_scores=0.7514816,
final_local_scores=0.7765861, pem=-0.2135508,
similarity=28.425913]
17: 3930485 Andrei_Olhovskiy
    [pem_pos=0, final_scores=0.82411236,
final_local_scores=0.92857665, pem=0.0, similarity=31.800293]
18: 25391 Russia
    [pem_pos=0, final_scores=0.89667886,
final_local_scores=0.96554667, pem=-0.20299491,
similarity=34.00747]
19: 1918850 Mark_Woodforde
    [pem_pos=0, final_scores=0.8548092,
final_local_scores=0.96307355, pem=0.0, similarity=32.82975]
20: 4689264 Australia
```

```
[pem_pos=0, final_scores=0.77874357,
final_local_scores=0.8252157, pem=-0.17168477,
similarity=29.649445]
21: 261368 Mark_Philippoussis
    [pem_pos=0, final_scores=0.93544346,
final_local_scores=1.0100493, pem=0.0, similarity=34.2316]
22: 4689264 Australia
    [pem_pos=0, final_scores=0.78512734,
final_local_scores=0.83218896, pem=-0.17168477,
similarity=29.85754]
23: 1723684 Roberto_Carretero
    [pem_pos=0, final_scores=0.5802719,
final_local_scores=0.6903825, pem=0.0, similarity=24.692102]
24: 26667 Spain
    [pem_pos=0, final_scores=0.6328872,
final_local_scores=0.6445703, pem=-0.18857051,
similarity=24.350452]
25: 14891740 Jordi_Burillo
    [pem_pos=0, final_scores=0.64410335,
final_local_scores=0.7607378, pem=0.0, similarity=26.791645]
26: 26667 Spain
    [pem_pos=0, final_scores=0.58888143,
final_local_scores=0.5965007, pem=-0.18857051,
similarity=22.915958]
27: 4076952 Francisco_Clavet
    [pem_pos=0, final_scores=0.6487657,
final_local_scores=0.75253254, pem=0.0, similarity=26.546785]
28: 26667 Spain
    [pem_pos=0, final_scores=0.60531855,
final_local_scores=0.61445576, pem=-0.18857051,
similarity=23.451775]
29: 979076 Cédric_Pioline
    [pem_pos=0, final_scores=0.8359646,
final_local_scores=0.91387016, pem=0.0, similarity=31.361422]
30: 5843419 France
    [pem_pos=0, final_scores=0.6845164,
final_local_scores=0.6968895, pem=-0.16240834,
similarity=25.769487]
31: 1723516 Wayne_Ferreira
    [pem_pos=0, final_scores=0.82309157,
final_local_scores=0.9212279, pem=0.0, similarity=31.58099]
32: 17416221 South_Africa
    [pem_pos=0, final_scores=0.62030464,
final_local_scores=0.67620695, pem=-0.26251382,
similarity=25.696669]
33: 3039956 Karol_Kučera
    [pem_pos=0, final_scores=0.77816516,
final_local_scores=0.8595281, pem=0.0, similarity=29.739746]
34: 26830 Slovakia
    [pem_pos=0, final_scores=0.673117,
final_local_scores=0.68576545, pem=-0.36602598,
similarity=26.544828]
```

35: 1964402 Jonas_Björkman
[pem_pos=0, final_scores=0.701813,
final_local_scores=0.7910792, pem=0.0, similarity=27.697094]
36: 5058739 Sweden
[pem_pos=0, final_scores=0.7667971,
final_local_scores=0.79331607, pem=-0.2135508,
similarity=28.925169]
37: 21241 Norway
[pem_pos=0, final_scores=0.7626712,
final_local_scores=0.82502425, pem=-0.22668563,
similarity=29.942833]
38: 17733765 Alex_Corretja
[pem_pos=0, final_scores=0.8809772,
final_local_scores=0.9466668, pem=0.0, similarity=32.340137]
39: 26667 Spain
[pem_pos=0, final_scores=0.6620144,
final_local_scores=0.6763874, pem=-0.18857051,
similarity=25.299938]
40: 1869714 Byron_Black
[pem_pos=0, final_scores=0.72117233,
final_local_scores=0.81656915, pem=0.0, similarity=28.457766]
41: 34399 Zimbabwe
[pem_pos=0, final_scores=0.4610774,
final_local_scores=0.5229088, pem=-0.36737105,
similarity=21.692175]
42: 10250905 David_Rikl
[pem_pos=0, final_scores=0.72174704,
final_local_scores=0.8239328, pem=0.0, similarity=28.677511]
43: 5321 Czech_Republic
[pem_pos=0, final_scores=0.5570202,
final_local_scores=0.56038773, pem=-0.27794555,
similarity=22.324312]
44: 2508144 Hicham_Arazi
[pem_pos=0, final_scores=0.7073393,
final_local_scores=0.80602306, pem=0.0, similarity=28.14305]
45: 19291 Morocco
[pem_pos=0, final_scores=0.61694884,
final_local_scores=0.65971565, pem=-0.29598162,
similarity=25.386538]
46: 2493300 Sjeng_Schalken
[pem_pos=0, final_scores=0.7829152,
final_local_scores=0.8649036, pem=0.0, similarity=29.900164]
47: 21148 Netherlands
[pem_pos=0, final_scores=0.6155832,
final_local_scores=0.62409073, pem=-0.24940795,
similarity=24.070145]
48: 20856583 Gilbert_Schaller
[pem_pos=0, final_scores=0.926894,
final_local_scores=1.0529342, pem=0.0, similarity=35.51137]
49: 26964606 Austria
[pem_pos=0, final_scores=0.6978465,
final_local_scores=0.71120214, pem=-0.25262418,

```
similarity=26.687214]
50: 4026564 Grant_Stafford
    [pem_pos=0, final_scores=0.81510776,
final_local_scores=0.928747, pem=0.0, similarity=31.805374]
51: 17416221 South_Africa
    [pem_pos=0, final_scores=0.60910666,
final_local_scores=0.6639748, pem=-0.26251382,
similarity=25.331636]
52: 1723582 Guy_Forget
    [pem_pos=0, final_scores=0.7443506,
final_local_scores=0.82064646, pem=0.0, similarity=28.579441]
53: 5843419 France
    [pem_pos=0, final_scores=0.741561,
final_local_scores=0.7592021, pem=-0.16240834,
similarity=27.629019]
54: 2989127 Fernando_Meligeni
    [pem_pos=0, final_scores=0.5721416,
final_local_scores=0.61269736, pem=0.0, similarity=22.373825]
55: 3383 Brazil
    [pem_pos=0, final_scores=0.78395444,
final_local_scores=0.8224746, pem=-0.21313342,
similarity=29.793047]
56: 489963 Yevgeny_Kafelnikov
    [pem_pos=0, final_scores=0.70231485,
final_local_scores=0.7694422, pem=0.0, similarity=27.051403]
57: 25391 Russia
    [pem_pos=0, final_scores=0.87199646,
final_local_scores=0.93858486, pem=-0.20299491,
similarity=33.202873]
58: 194244 Goran_Ivanišević
    [pem_pos=0, final_scores=1.0091091,
final_local_scores=1.0898104, pem=0.0, similarity=36.611828]
59: 5573 Croatia
    [pem_pos=0, final_scores=0.6294928,
final_local_scores=0.64846814, pem=-0.35021794,
similarity=25.345835]
60: 1723532 Andrei_Chесnokов
    [pem_pos=0, final_scores=0.75216883,
final_local_scores=0.84427917, pem=0.0, similarity=29.284687]
61: 25391 Russia
    [pem_pos=0, final_scores=0.8542798,
final_local_scores=0.9192321, pem=-0.20299491,
similarity=32.625347]
62: 3983307 Scott_Draper
    [pem_pos=0, final_scores=0.72186273,
final_local_scores=0.83507663, pem=0.0, similarity=29.010065]
63: 4689264 Australia
    [pem_pos=0, final_scores=0.7247413,
final_local_scores=0.7662264, pem=-0.17168477,
similarity=27.889086]
64: 441573 Galo_Blanco
    [pem_pos=0, final_scores=0.6754504,
```

```
final_local_scores=0.7846676, pem=0.0, similarity=27.50576]
65: 26667 Spain
    [pem_pos=0, final_scores=0.6504157,
final_local_scores=0.66371757, pem=-0.18857051,
similarity=24.921844]
66: 3929800 Renzo_Furlan
    [pem_pos=0, final_scores=0.75160867,
final_local_scores=0.85472476, pem=0.0, similarity=29.596405]
67: 14532 Italy
    [pem_pos=0, final_scores=0.7086301,
final_local_scores=0.7555619, pem=-0.20612453,
similarity=27.758125]
68: 1408278 Thomas_Johansson
    [pem_pos=0, final_scores=0.75749665,
final_local_scores=0.82678646, pem=0.0, similarity=28.76267]
69: 5058739 Sweden
    [pem_pos=0, final_scores=0.7456887,
final_local_scores=0.77025825, pem=-0.2135508,
similarity=28.237078]
70: 11867 Germany
    [pem_pos=0, final_scores=0.8405902,
final_local_scores=0.87154025, pem=-0.15529114,
similarity=30.94271]
71: 517430 Greg_Rusedski
    [pem_pos=0, final_scores=1.0085857,
final_local_scores=1.0838976, pem=0.0, similarity=36.43538]
72: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.57342714,
final_local_scores=0.6037875, pem=-0.61114657,
similarity=25.431446]
73: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=0.8551298,
final_local_scores=0.9369784, pem=0.0, similarity=32.051018]
74: 31750 Ukraine
    [pem_pos=0, final_scores=0.568361,
final_local_scores=0.5958655, pem=-0.35407752,
similarity=23.797056]
75: 4026812 Jean-Philippe_Fleurian
    [pem_pos=0, final_scores=0.6589517,
final_local_scores=0.7564946, pem=0.0, similarity=26.66502]
76: 5843419 France
    [pem_pos=0, final_scores=0.74465805,
final_local_scores=0.7625852, pem=-0.16240834,
similarity=27.729979]
77: 3930155 Ján_Krošlák
    [pem_pos=0, final_scores=0.69508433,
final_local_scores=0.77093, pem=0.0, similarity=27.095802]
78: 26830 Slovakia
    [pem_pos=0, final_scores=0.67971134,
final_local_scores=0.6929688, pem=-0.36602598,
similarity=26.75979]
79: 2515285 Chris_Woodruff
```

[pem_pos=0, final_scores=0.66748446,
final_local_scores=0.77054304, pem=0.0, similarity=27.084253]
80: 3434750 United_States
[pem_pos=0, final_scores=0.5047625,
final_local_scores=0.5088678, pem=-0.14075686,
similarity=20.0408]
81: 1439769 Petr_Korda
[pem_pos=0, final_scores=0.94977146,
final_local_scores=1.050184, pem=0.0, similarity=35.4293]
82: 5321 Czech_Republic
[pem_pos=0, final_scores=0.5545209,
final_local_scores=0.55765766, pem=-0.27794555,
similarity=22.242842]
83: 2805841 Bohdan_Ulihrach
[pem_pos=0, final_scores=0.8135353,
final_local_scores=0.92981535, pem=0.0, similarity=31.837257]
84: 5321 Czech_Republic
[pem_pos=0, final_scores=0.6382666,
final_local_scores=0.64913714, pem=-0.27794555,
similarity=24.97277]
85: 26667 Spain
[pem_pos=0, final_scores=0.5736265,
final_local_scores=0.579837, pem=-0.18857051,
similarity=22.418682]
86: 871725 Todd_Martin
[pem_pos=0, final_scores=0.86638683,
final_local_scores=0.94103175, pem=0.0, similarity=32.171978]
87: 3434750 United_States
[pem_pos=0, final_scores=0.50056285,
final_local_scores=0.5042804, pem=-0.14075686,
similarity=19.903902]
88: 19291 Morocco
[pem_pos=0, final_scores=0.65388596,
final_local_scores=0.7000638, pem=-0.29598162,
similarity=26.590607]
89: 3929637 Andrea_Gaudenzi
[pem_pos=0, final_scores=0.8569748,
final_local_scores=0.9739906, pem=0.0, similarity=33.155537]
90: 14532 Italy
[pem_pos=0, final_scores=0.7189695,
final_local_scores=0.76685613, pem=-0.20612453,
similarity=28.095167]
91: 2575718 Shuzo_Matsuoka
[pem_pos=0, final_scores=0.67439014,
final_local_scores=0.79434663, pem=0.0, similarity=27.794601]
92: 15573 Japan
[pem_pos=0, final_scores=0.73576003,
final_local_scores=0.7961918, pem=-0.13549899,
similarity=28.586529]
93: 4025956 Doug_Flach
[pem_pos=0, final_scores=0.5132647,
final_local_scores=0.61798275, pem=0.0, similarity=22.531551]

94: 3434750 United_States
[pem_pos=0, final_scores=0.46493578,
final_local_scores=0.4653632, pem=-0.14075686,
similarity=18.742537]
95: 200446 Mats_Wilander
[pem_pos=0, final_scores=0.9299714,
final_local_scores=0.99949664, pem=0.0, similarity=33.916687]
96: 5058739 Sweden
[pem_pos=0, final_scores=0.80114025,
final_local_scores=0.8308307, pem=-0.2135508,
similarity=30.04468]
97: 262030 Tim_Henman
[pem_pos=0, final_scores=0.79084927,
final_local_scores=0.8330632, pem=0.0, similarity=28.94998]
98: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5902558,
final_local_scores=0.62217015, pem=-0.61114657,
similarity=25.980022]
99: 2186352 Paul_Haarhuis
[pem_pos=0, final_scores=0.6218633,
final_local_scores=0.7145324, pem=0.0, similarity=25.412785]
100: 21148 Netherlands
[pem_pos=0, final_scores=0.6569108,
final_local_scores=0.6692349, pem=-0.24940795,
similarity=25.417336]
101: 4026170 Michael_Joyce_(tennis)
[pem_pos=1, final_scores=0.50900733,
final_local_scores=0.59406424, pem=-0.41227743,
similarity=24.059803]
102: 3434750 United_States
[pem_pos=0, final_scores=0.5004479,
final_local_scores=0.50415486, pem=-0.14075686,
similarity=19.900158]
103: 13355036 Michael_Tebbutt
[pem_pos=0, final_scores=0.6068327,
final_local_scores=0.72405803, pem=0.0, similarity=25.697048]
104: 4689264 Australia
[pem_pos=0, final_scores=0.7292592,
final_local_scores=0.77116156, pem=-0.17168477,
similarity=28.036362]
105: 871728 Richey_Reneberg
[pem_pos=0, final_scores=0.8727086,
final_local_scores=0.98060864, pem=0.0, similarity=33.35303]
106: 3434750 United_States
[pem_pos=0, final_scores=0.5189093,
final_local_scores=0.524321, pem=-0.14075686,
similarity=20.501957]
107: 3079144 Jonathan_Stark_(tennis)
[pem_pos=1, final_scores=0.59481084,
final_local_scores=0.67397094, pem=-0.86421156,
similarity=28.902067]
108: 3434750 United_States

[pem_pos=0, final_scores=0.49090147,
final_local_scores=0.4937268, pem=-0.14075686,
similarity=19.588964]
109: 15290626 Bernd_Karbacher
[pem_pos=0, final_scores=0.62546784,
final_local_scores=0.7070299, pem=0.0, similarity=25.188894]
110: 11867 Germany
[pem_pos=0, final_scores=0.89769274,
final_local_scores=0.93391603, pem=-0.15529114,
similarity=32.804127]
111: 391932 Stefan_Edberg
[pem_pos=0, final_scores=0.87583655,
final_local_scores=0.93872076, pem=0.0, similarity=32.103012]
112: 5058739 Sweden
[pem_pos=0, final_scores=0.7730163,
final_local_scores=0.80010957, pem=-0.2135508,
similarity=29.127901]
113: 161972 Richard_Krajicek
[pem_pos=0, final_scores=0.9957331,
final_local_scores=1.0855303, pem=0.0, similarity=36.484104]
114: 21148 Netherlands
[pem_pos=0, final_scores=0.6014517,
final_local_scores=0.6086542, pem=-0.24940795,
similarity=23.609488]
115: 595 Andre_Agassi
[pem_pos=0, final_scores=0.9014825,
final_local_scores=0.94717795, pem=0.0, similarity=32.355392]
116: 3434750 United_States
[pem_pos=0, final_scores=0.50846726,
final_local_scores=0.5129147, pem=-0.14075686,
similarity=20.16157]
117: 20694823 Mauricio_Hadad
[pem_pos=0, final_scores=0.61869895,
final_local_scores=0.7245488, pem=0.0, similarity=25.711695]
118: 5222 Colombia
[pem_pos=0, final_scores=0.65104157,
final_local_scores=0.684667, pem=-0.22328302,
similarity=25.73579]
119: 3930413 Marcos_Ondruska
[pem_pos=0, final_scores=0.93586963,
final_local_scores=1.0380616, pem=0.0, similarity=35.06754]
120: 17416221 South_Africa
[pem_pos=0, final_scores=0.6170164,
final_local_scores=0.672615, pem=-0.26251382,
similarity=25.589478]
121: 1723559 Félix_Mantilla_Botella
[pem_pos=0, final_scores=0.77564293,
final_local_scores=0.8505147, pem=0.0, similarity=29.470768]
122: 26667 Spain
[pem_pos=0, final_scores=0.6319016,
final_local_scores=0.6434937, pem=-0.18857051,
similarity=24.318325]

123: 669421 Carlos_Moyá
[pem_pos=0, final_scores=0.8648817,
final_local_scores=0.9370896, pem=0.0, similarity=32.054337]
124: 26667 Spain
[pem_pos=0, final_scores=0.62692326,
final_local_scores=0.6380556, pem=-0.18857051,
similarity=24.15604]
125: 20867776 Scott_Humphries
[pem_pos=0, final_scores=0.70532507,
final_local_scores=0.81073606, pem=0.0, similarity=28.283693]
126: 3434750 United_States
[pem_pos=0, final_scores=0.5171353,
final_local_scores=0.5223833, pem=-0.14075686,
similarity=20.44413]
127: 2876758 Jan_Siemerink
[pem_pos=0, final_scores=0.7450629,
final_local_scores=0.8327866, pem=0.0, similarity=28.941727]
128: 21148 Netherlands
[pem_pos=0, final_scores=0.6399996,
final_local_scores=0.65076196, pem=-0.24940795,
similarity=24.866068]
129: 3254714 Carl-Uwe_Staab
[pem_pos=0, final_scores=0.65947974,
final_local_scores=0.75943184, pem=0.0, similarity=26.752674]
130: 11867 Germany
[pem_pos=0, final_scores=0.8709981,
final_local_scores=0.90475625, pem=-0.15529114,
similarity=31.933943]
131: 1829027 David_Wheaton
[pem_pos=0, final_scores=0.55607355,
final_local_scores=0.6190195, pem=0.0, similarity=22.56249]
132: 3434750 United_States
[pem_pos=0, final_scores=0.48645854,
final_local_scores=0.48887354, pem=-0.14075686,
similarity=19.444132]
133: 2075801 Kevin_Kim
[pem_pos=0, final_scores=0.46892798,
final_local_scores=0.55076206, pem=0.0, similarity=20.52555]
134: 3434750 United_States
[pem_pos=0, final_scores=0.4911253,
final_local_scores=0.4939713, pem=-0.14075686,
similarity=19.59626]
135: 2817661 Nicolás_Lapentti
[pem_pos=0, final_scores=0.79730636,
final_local_scores=0.85835695, pem=0.0, similarity=29.704796]
136: 9334 Ecuador
[pem_pos=0, final_scores=0.48115993,
final_local_scores=0.5432853, pem=-0.21404564,
similarity=21.466444]
137: 3434750 United_States
[pem_pos=0, final_scores=0.5019874,
final_local_scores=0.5058364, pem=-0.14075686,

similarity=19.95034]
138: 3930942 Karim_Alami
[pem_pos=0, final_scores=0.64499825,
final_local_scores=0.7450731, pem=0.0, similarity=26.32418]
139: 19291 Morocco
[pem_pos=0, final_scores=0.6344953,
final_local_scores=0.6788825, pem=-0.29598162,
similarity=25.958513]
140: 873616 MaliVai_Washington
[pem_pos=0, final_scores=0.7851326,
final_local_scores=0.8517895, pem=0.0, similarity=29.50881]
141: 3434750 United_States
[pem_pos=0, final_scores=0.49432898,
final_local_scores=0.4974708, pem=-0.14075686,
similarity=19.700693]
142: 933709 Thomas_Enqvist
[pem_pos=0, final_scores=0.8927639,
final_local_scores=0.9701734, pem=0.0, similarity=33.041622]
143: 5058739 Sweden
[pem_pos=0, final_scores=0.7847975,
final_local_scores=0.81297874, pem=-0.2135508,
similarity=29.511942]
144: 22577800 Stéphane_Simian
[pem_pos=0, final_scores=0.7487113,
final_local_scores=0.8829729, pem=0.0, similarity=30.439386]
145: 5843419 France
[pem_pos=0, final_scores=0.7342182,
final_local_scores=0.75118124, pem=-0.16240834,
similarity=27.38966]
146: 3930324 Guillaume_Raoux
[pem_pos=0, final_scores=0.73274416,
final_local_scores=0.821522, pem=0.0, similarity=28.605568]
147: 5843419 France
[pem_pos=0, final_scores=0.722722,
final_local_scores=0.7386234, pem=-0.16240834,
similarity=27.01491]
148: 3004535 Filip_Dewulf
[pem_pos=0, final_scores=0.75115913,
final_local_scores=0.8313651, pem=0.0, similarity=28.899305]
149: 3343 Belgium
[pem_pos=0, final_scores=0.6989789,
final_local_scores=0.70167327, pem=-0.24180195,
similarity=26.344]
150: 2261491 Mark_Knowles
[pem_pos=0, final_scores=0.57615024,
final_local_scores=0.6567282, pem=0.0, similarity=23.687792]
151: 3451 The_Bahamas
[pem_pos=0, final_scores=0.4625979,
final_local_scores=0.5034777, pem=-0.35113433,
similarity=21.024015]
152: 2936118 Marcelo_Filippini
[pem_pos=0, final_scores=0.8032399,

```
final_local_scores=0.8932039, pem=0.0, similarity=30.7447]
153: 31843 Uruguay
    [pem_pos=0, final_scores=0.48151487,
final_local_scores=0.48833102, pem=-0.44594222,
similarity=21.087587]
154: 1547951 Todd_Woodbridge
    [pem_pos=0, final_scores=0.81538886,
final_local_scores=0.9037114, pem=0.0, similarity=31.058264]
155: 4689264 Australia
    [pem_pos=0, final_scores=0.73658985,
final_local_scores=0.77916914, pem=-0.17168477,
similarity=28.275324]
156: 3343 Belgium
    [pem_pos=0, final_scores=0.7007824,
final_local_scores=0.7036433, pem=-0.24180195,
similarity=26.40279]
157: 913625 Sergi_Bruguera
    [pem_pos=0, final_scores=0.95262176,
final_local_scores=1.0353656, pem=0.0, similarity=34.987087]
158: 26667 Spain
    [pem_pos=0, final_scores=0.62241596,
final_local_scores=0.6331321, pem=-0.18857051,
similarity=24.009113]
159: 979088 Michael_Stich
    [pem_pos=0, final_scores=0.8272111,
final_local_scores=0.9131705, pem=0.0, similarity=31.340542]
160: 11867 Germany
    [pem_pos=0, final_scores=0.8729113,
final_local_scores=0.9068461, pem=-0.15529114,
similarity=31.996307]
161: 20930737 Chuck_Adams
    [pem_pos=0, final_scores=0.47202468,
final_local_scores=0.5538646, pem=0.0, similarity=20.618137]
162: 3434750 United_States
    [pem_pos=0, final_scores=0.52550167,
final_local_scores=0.5315222, pem=-0.14075686,
similarity=20.716854]
163: 3930198 Javier_Frana
    [pem_pos=0, final_scores=0.5761125,
final_local_scores=0.645222, pem=0.0, similarity=23.344425]
164: 18951905 Argentina
    [pem_pos=0, final_scores=0.6371284,
final_local_scores=0.6527497, pem=-0.30003807,
similarity=25.20072]
165: 166170 Thomas_Muster
    [pem_pos=0, final_scores=0.9540413,
final_local_scores=1.0196857, pem=0.0, similarity=34.519173]
166: 26964606 Austria
    [pem_pos=0, final_scores=0.6418201,
final_local_scores=0.6500018, pem=-0.25262418,
similarity=24.860874]
167: 685780 Jim_Courier
```

[pem_pos=0, final_scores=0.90264434,
final_local_scores=0.96590155, pem=0.0, similarity=32.914143]
168: 3434750 United_States
[pem_pos=0, final_scores=0.51306015,
final_local_scores=0.51793176, pem=-0.14075686,
similarity=20.311289]
169: 420527 Javier_Sánchez
[pem_pos=0, final_scores=0.56494504,
final_local_scores=0.6398468, pem=0.0, similarity=23.184017]
170: 26667 Spain
[pem_pos=0, final_scores=0.6025127,
final_local_scores=0.6113908, pem=-0.18857051,
similarity=23.36031]
171: 1915197 Jim_Grabb
[pem_pos=0, final_scores=0.81614965,
final_local_scores=0.91834146, pem=0.0, similarity=31.494854]
172: 3434750 United_States
[pem_pos=0, final_scores=0.49627173,
final_local_scores=0.49959296, pem=-0.14075686,
similarity=19.76402]
173: 3823337 Sandon_Stolle
[pem_pos=0, final_scores=0.86592966,
final_local_scores=0.9681786, pem=0.0, similarity=32.982094]
174: 4689264 Australia
[pem_pos=0, final_scores=0.8206412,
final_local_scores=0.8709826, pem=-0.17168477,
similarity=31.015217]
175: 23297 Patrick_Rafter
[pem_pos=0, final_scores=0.98462564,
final_local_scores=1.0673597, pem=0.0, similarity=35.941856]
176: 4689264 Australia
[pem_pos=0, final_scores=0.67863137,
final_local_scores=0.7158583, pem=-0.17168477,
similarity=26.386002]
177: 1804967 Kenneth_Carlsen
[pem_pos=0, final_scores=0.8005188,
final_local_scores=0.8908344, pem=0.0, similarity=30.673986]
178: 76972 Denmark
[pem_pos=0, final_scores=0.5938531,
final_local_scores=0.61797696, pem=-0.26762357,
similarity=23.986757]
179: 2598440 Jason_Stoltenberg
[pem_pos=0, final_scores=0.7769117,
final_local_scores=0.8549581, pem=0.0, similarity=29.603369]
180: 4689264 Australia
[pem_pos=0, final_scores=0.7096519,
final_local_scores=0.7497436, pem=-0.17168477,
similarity=27.397205]
181: 3930788 Stefano_Pescosolido
[pem_pos=0, final_scores=0.8310861,
final_local_scores=0.94623977, pem=0.0, similarity=32.327396]
182: 14532 Italy

[pem_pos=0, final_scores=0.73506767,
final_local_scores=0.7844409, pem=-0.20612453,
similarity=28.61993]
183: 3930278 Arnaud_Boetsch
[pem_pos=0, final_scores=0.6339287,
final_local_scores=0.7284755, pem=0.0, similarity=25.828875]
184: 5843419 France
[pem_pos=0, final_scores=0.73375493,
final_local_scores=0.75067526, pem=-0.16240834,
similarity=27.374561]
185: 3929855 Nicolás_Pereira
[pem_pos=0, final_scores=0.6239586,
final_local_scores=0.6950262, pem=0.0, similarity=24.83068]
186: 32374 Venezuela
[pem_pos=0, final_scores=0.48080707,
final_local_scores=0.49820745, pem=-0.21226841,
similarity=20.111567]
187: 2785107 Carlos_Costa_(tennis)
[pem_pos=0, final_scores=0.42445183,
final_local_scores=0.47099805, pem=-0.021202235,
similarity=18.260534]
188: 26667 Spain
[pem_pos=0, final_scores=0.60221183,
final_local_scores=0.61106217, pem=-0.18857051,
similarity=23.350502]
189: 2793059 Magnus_Gustafsson
[pem_pos=0, final_scores=0.745285,
final_local_scores=0.8136529, pem=0.0, similarity=28.370737]
190: 5058739 Sweden
[pem_pos=0, final_scores=0.7546504,
final_local_scores=0.78004754, pem=-0.2135508,
similarity=28.52921]
191: 1862751 Jeff_Tarango
[pem_pos=0, final_scores=0.68916404,
final_local_scores=0.77320874, pem=0.0, similarity=27.163803]
192: 3434750 United_States
[pem_pos=0, final_scores=0.5117751,
final_local_scores=0.51652807, pem=-0.14075686,
similarity=20.269398]
193: 4026041 Alex_Rădulescu
[pem_pos=0, final_scores=0.5753266,
final_local_scores=0.6931949, pem=0.0, similarity=24.776031]
194: 11867 Germany
[pem_pos=0, final_scores=0.84586245,
final_local_scores=0.87729937, pem=-0.15529114,
similarity=31.114574]
195: 483554 Marcelo_Ríos
[pem_pos=0, final_scores=0.66046983,
final_local_scores=0.70344096, pem=0.0, similarity=25.081795]
196: 5489 Chile
[pem_pos=0, final_scores=0.60294986,
final_local_scores=0.6409239, pem=-0.27404457,

similarity=24.706457]
197: 1717291 Marc_Rosset
[pem_pos=0, final_scores=0.80474883,
final_local_scores=0.89762217, pem=0.0, similarity=30.876549]
198: 26748 Switzerland
[pem_pos=0, final_scores=0.6826368,
final_local_scores=0.69294226, pem=-0.18598008,
similarity=25.77988]
199: 3056059 Jared_Palmer
[pem_pos=0, final_scores=0.66714627,
final_local_scores=0.7736189, pem=0.0, similarity=27.176043]
200: 3434750 United_States
[pem_pos=0, final_scores=0.51502186,
final_local_scores=0.52007467, pem=-0.14075686,
similarity=20.375235]
201: 4026614 Martin_Damm
[pem_pos=0, final_scores=0.78807205,
final_local_scores=0.8776377, pem=0.0, similarity=30.28017]
202: 5321 Czech_Republic
[pem_pos=0, final_scores=0.5593003,
final_local_scores=0.56287843, pem=-0.27794555,
similarity=22.39864]
203: 3929963 Hernán_Gumy
[pem_pos=0, final_scores=0.64831877,
final_local_scores=0.7400878, pem=0.0, similarity=26.17541]
204: 18951905 Argentina
[pem_pos=0, final_scores=0.5467649,
final_local_scores=0.5540413, pem=-0.30003807,
similarity=22.255064]
205: 2987792 Nicklas_Kulti
[pem_pos=0, final_scores=0.7816382,
final_local_scores=0.8833615, pem=0.0, similarity=30.450981]
206: 5058739 Sweden
[pem_pos=0, final_scores=0.78948313,
final_local_scores=0.818097, pem=-0.2135508,
similarity=29.66468]
207: 1727962 Jakob_Hlasek
[pem_pos=0, final_scores=0.97541744,
final_local_scores=1.0919356, pem=0.0, similarity=36.675247]
208: 26748 Switzerland
[pem_pos=0, final_scores=0.68301713,
final_local_scores=0.69335777, pem=-0.18598008,
similarity=25.79228]
209: 10980025 Cecil_Mamait
[pem_pos=0, final_scores=0.7220583,
final_local_scores=0.82685125, pem=0.0, similarity=28.764603]
210: 3434750 United_States
[pem_pos=0, final_scores=0.4999978,
final_local_scores=0.5036631, pem=-0.14075686,
similarity=19.885483]
211: 1728009 Alberto_Berasategui
[pem_pos=0, final_scores=0.7141741,

```
final_local_scores=0.7823642, pem=0.0, similarity=27.437021]
212: 26667 Spain
    [pem_pos=0, final_scores=0.63061553,
final_local_scores=0.6420889, pem=-0.18857051,
similarity=24.276402]
213: 1574767 Vince_Spadea
    [pem_pos=0, final_scores=0.83663267,
final_local_scores=0.9122792, pem=0.0, similarity=31.313942]
214: 3434750 United_States
    [pem_pos=0, final_scores=0.5032932,
final_local_scores=0.5072629, pem=-0.14075686,
similarity=19.992908]
215: 3930665 Daniel_Vacek
    [pem_pos=0, final_scores=0.8499597,
final_local_scores=0.9659032, pem=0.0, similarity=32.914192]
216: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.55170184,
final_local_scores=0.55457824, pem=-0.27794555,
similarity=22.150946]
217: 3930698 David_Prinosil
    [pem_pos=0, final_scores=0.80733746,
final_local_scores=0.9040528, pem=0.0, similarity=31.06845]
218: 11867 Germany
    [pem_pos=0, final_scores=0.8349667,
final_local_scores=0.8653974, pem=-0.15529114,
similarity=30.759396]
219: 13833030 Tomás_Carbonell
    [pem_pos=0, final_scores=0.71151006,
final_local_scores=0.81793964, pem=0.0, similarity=28.498663]
220: 26667 Spain
    [pem_pos=0, final_scores=0.630479,
final_local_scores=0.6419397, pem=-0.18857051,
similarity=24.27195]
221: 391834 Michael_Chang
    [pem_pos=0, final_scores=0.77528554,
final_local_scores=0.82039946, pem=0.0, similarity=28.57207]
222: 3434750 United_States
    [pem_pos=0, final_scores=0.5003448,
final_local_scores=0.5040422, pem=-0.14075686,
similarity=19.896795]
```

Sample: b'50_BASEBALL'

```
BASEBALL - [1 1 ORIOLES 1] ' MANAGER [2 2 DAVEY JOHNSON 2]
HOSPITALIZED . [3 3 BALTIMORE 3] 1996-08-22 [4 4 Baltimore Or
ioles 4] manager [5 5 Davey Johnson 5] will miss Thursday
night 's game against the [6 6 Seattle Mariners 6] after
being admitted to a hospital with an irregular heartbeat .
The 53-year-old [7 7 Johnson 7] was hospitalized after
experiencing dizziness . " He is in no danger and will be
treated and observed this evening , " said [8 8 Orioles 8]
team physician Dr. William Goldiner , adding that [9 9 Johnson
9] is expected to be released on Friday . [10 10 Orioles 10]
```

' bench coach [11 11 Andy Etchebarren 11] will manage the club in [12 12 Johnson 12] 's absence . [13 13 Johnson 13] is the second manager to be hospitalized this week after [14 14 California Angels 14] skipper [15 15 John McNamara 15] was admitted to [16 16 New York 16] 's [17 17 Columbia Presbyterian Hospital 17] on Wednesday with a blood clot in his left calf . [18 18 Johnson 18] , who played eight seasons in [19 19 Baltimore 19] , was named [20 20 Orioles 20] manager in the off-season replacing [21 21 Phil Regan 21] . He led the [22 22 Cincinnati Reds 22] to the National League Championship Series last year and guided the [23 23 New York Mets 23] to a [24 24 World Series 24] championship in 1986 . [25 25 Baltimore 25] has won 16 of its last 22 games to pull within five games of the slumping [26 26 New York Yankees 26] in the [27 27 American League East Division 27] .

1: 4340 Baltimore_Orioles
2: 652850 Davey_Johnson
3: 26997138 Baltimore
4: 4340 Baltimore_Orioles
5: 652850 Davey_Johnson
6: 27660 Seattle_Mariners
7: 652850 Davey_Johnson
8: 4340 Baltimore_Orioles
9: 652850 Davey_Johnson
10: 4340 Baltimore_Orioles
11: 1928809 Andy_Etchebarren
12: 652850 Davey_Johnson
13: 652850 Davey_Johnson
14: 1360083 Los_Angeles_Angels_of_Anheim
15: 2226049 John_McNamara_(baseball)
16: 645042 New_York_City
17: 1035828 NewYork-Presbyterian_Hospital
18: 652850 Davey_Johnson
19: 4340 Baltimore_Orioles
20: 4340 Baltimore_Orioles
21: 1731893 Phil_Regan_(baseball)
22: 6671 Cincinnati_Reds
23: 21728 New_York_Mets
24: 23489161 World_Series
25: 4340 Baltimore_Orioles
26: 4848143 New_York_Yankees
27: 2004582 American_League_East
1: 4340 Baltimore_Orioles
[pem_pos=0, final_scores=0.5437277,
final_local_scores=0.48492318, pem=-0.18171716,
similarity=19.548994]
2: 652850 Davey_Johnson
[pem_pos=0, final_scores=0.46837074,
final_local_scores=0.43166023, pem=0.0, similarity=16.971315]
3: 26997138 Baltimore
[pem_pos=0, final_scores=0.52648884,
final_local_scores=0.5913671, pem=-0.19949137,

```
similarity=22.822153]
4: 4340 Baltimore_Orioles
    [pem_pos=0, final_scores=0.44548458,
final_local_scores=0.37760758, pem=-0.21811225,
similarity=16.544405]
5: 652850 Davey_Johnson
    [pem_pos=0, final_scores=0.604307,
final_local_scores=0.5801499, pem=0.0, similarity=21.402542]
6: 27660 Seattle_Mariners
    [pem_pos=0, final_scores=0.45113713,
final_local_scores=0.3990531, pem=-0.008242749,
similarity=16.04308]
7: 652850 Davey_Johnson
    [pem_pos=0, final_scores=0.70240957,
final_local_scores=0.68731207, pem=0.0, similarity=24.600475]
8: 4340 Baltimore_Orioles
    [pem_pos=0, final_scores=0.5026378,
final_local_scores=0.4400388, pem=-0.18171716,
similarity=18.209555]
9: 652850 Davey_Johnson
    [pem_pos=0, final_scores=0.67276555,
final_local_scores=0.6549305, pem=0.0, similarity=23.634144]
10: 4340 Baltimore_Orioles
    [pem_pos=0, final_scores=0.462824,
final_local_scores=0.3965482, pem=-0.18171716,
similarity=16.911709]
11: 1928809 Andy_Etchebarren
    [pem_pos=0, final_scores=0.6598024,
final_local_scores=0.72427607, pem=0.0, similarity=25.703556]
12: 652850 Davey_Johnson
    [pem_pos=0, final_scores=0.8534035,
final_local_scores=0.85225, pem=0.0, similarity=29.522552]
13: 652850 Davey_Johnson
    [pem_pos=0, final_scores=0.875575,
final_local_scores=0.8764691, pem=0.0, similarity=30.245298]
14: 1360083 Los_Angeles_Angels_ofAnaheim
    [pem_pos=0, final_scores=0.5315698,
final_local_scores=0.4821328, pem=-0.016987266,
similarity=18.569897]
15: 2226049 John_McNamara_(baseball)
    [pem_pos=0, final_scores=0.5497411,
final_local_scores=0.56208664, pem=-0.0888523,
similarity=21.346693]
16: 645042 New_York_City
    [pem_pos=1, final_scores=0.4101299,
final_local_scores=0.48201728, pem=-1.7906687,
similarity=28.212004]
17: 1035828 NewYork-Presbyterian_Hospital
    [pem_pos=0, final_scores=0.23626612,
final_local_scores=0.33325034, pem=0.0, similarity=14.034568]
18: 652850 Davey_Johnson
    [pem_pos=0, final_scores=0.9504662,
```

```

final_local_scores=0.95827633, pem=0.0, similarity=32.68659]
19: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=0.9140591,
final_local_scores=0.88945407, pem=-4.046561,
similarity=52.638622]
20: 4340 Baltimore_Orioles
    [pem_pos=0, final_scores=0.50469345,
final_local_scores=0.4422843, pem=-0.18171716,
similarity=18.276564]
21: 1731893 Phil_Regan_(baseball)
    [pem_pos=0, final_scores=0.6097315,
final_local_scores=0.6759081, pem=0.0, similarity=24.260159]
22: 6671 Cincinnati_Reds
    [pem_pos=0, final_scores=0.5772279,
final_local_scores=0.53265303, pem=-0.106723964,
similarity=20.565523]
23: 21728 New_York_Mets
    [pem_pos=0, final_scores=0.48296493,
final_local_scores=0.42459762, pem=-0.065782696,
similarity=17.11829]
24: 23489161 World_Series
    [pem_pos=0, final_scores=0.41607058,
final_local_scores=0.38033432, pem=-0.38314086,
similarity=17.523228]
25: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.0405197,
final_local_scores=1.0275931, pem=-4.046561,
similarity=56.76097]
26: 4848143 New_York_Yankees
    [pem_pos=0, final_scores=0.5909435,
final_local_scores=0.5420419, pem=-0.14222674,
similarity=21.038774]
27: 2004582 American_League_East
    [pem_pos=0, final_scores=0.3405823,
final_local_scores=0.3281564, pem=0.0, similarity=13.882555]

```

Sample: b'51_BASEBALL'

```

BASEBALL - [1 1 MAJOR LEAGUE 1] STANDINGS AFTER WEDNESDAY 'S
GAMES . [2 2 NEW YORK 2] 1996-08-22 [3 3 Major League Baseball
3] standings after games played on Wednesday ( tabulate under
won , lost , winning percentage and games behind ) : [4 4 AME
RICAN LEAGUE EASTERN DIVISION 4] W L PCT GB [5 5 NEW YORK 5]
72 53 .576 - [6 6 BALTIMORE 6] 67 58 .536 5 [7 7 BOSTON 7] 63
64 .496 10 [8 8 TORONTO 8] 58 69 .457 15 [9 9 DETROIT 9] 44
82 .349 28 1/2 [10 10 CENTRAL DIVISION 10] [11 11 CLEVELAND 11
] 76 51 .598 - 1 CHICAGO 69 59 .539 7 1/2 [12 12 MINNESOTA 12]
63 63 .500 12 1/2 [13 13 MILWAUKEE 13] 60 68 .469 16 1/2 [14 1
4 KANSAS CITY 14] 58 70 .453 18 1/2 [15 15 WESTERN DIVISION 15
] [16 16 TEXAS 16] 73 54 .575 - [17 17 SEATTLE 17] 64 61 .512
8 [18 18 OAKLAND 18] 62 67 .481 12 2 CALIFORNIA 58 68 .460 14
1/2 THURSDAY , AUGUST 22 SCHEDULE [19 19 OAKLAND 19] AT [20 20
BOSTON 20] [21 21 SEATTLE 21] AT [22 22 BALTIMORE 22] 3 CALIF

```

ORNIA AT [23 23 NEW YORK 23] [24 24 TORONTO 24] AT 4 CHICAGO [25 25 DETROIT 25] AT [26 26 KANSAS CITY 26] [27 27 TEXAS 27] AT [28 28 MINNESOTA 28] [29 29 NATIONAL LEAGUE EASTERN DIVISION 29] W L PCT GB [30 30 ATLANTA 30] 79 46 .632 - [31 31 MONTREAL 31] 67 58 .536 12 [32 32 NEW YORK 32] 59 69 .461 21 1/2 5 FLORIDA 58 69 .457 22 [33 33 PHILADELPHIA 33] 52 75 .409 28 [34 34 CENTRAL DIVISION 34] [35 35 HOUSTON 35] 68 59 .535 - ST LOUIS 67 59 .532 1/2 6 CHICAGO 63 62 .504 4 [36 36 CINCINNATI 36] 62 62 .500 4 1/2 [37 37 PITTSBURGH 37] 53 73 .421 14 1/2 [38 38 WESTERN DIVISION 38] [39 39 SAN DIEGO 39] 70 59 .543 - 7 LOS ANGELES 66 60 .524 2 1/2 [40 40 COLORADO 40] 65 62 .512 4 [41 41 SAN FRANCISCO 41] 54 70 .435 13 1/2 THURSDAY , AUGUST 22 SCHEDULE ST LOUIS AT [42 42 COLORADO 42] [43 43 CINCINNATI 43] AT [44 44 ATLANTA 44] [45 45 PITTSBURGH 45] AT [46 46 HOU STON 46] [47 47 PHILADELPHIA 47] AT 8 LOS ANGELES [48 48 MONTREAL 48] AT [49 49 SAN FRANCISCO 49]

1: 38776 Major_League_Baseball
2: 645042 New_York_City
3: 38776 Major_League_Baseball
4: 2004582 American_League_East
5: 4848143 New_York_Yankees
6: 4340 Baltimore_Orioles
7: 4339 Boston_Red_Sox
8: 30847 Toronto_Blue_Jays
9: 8579 Detroit_Tigers
10: 2456377 National_League_Central
11: 6652 Cleveland_Indians
12: 20050 Minnesota_Twins
13: 20152 Milwaukee_Brewers
14: 16883 Kansas_City_Royals
15: 2456474 National_League_West
16: 30857 Texas_Rangers_(baseball)
17: 27660 Seattle_Mariners
18: 22523 Oakland_Athletics
19: 22523 Oakland_Athletics
20: 4339 Boston_Red_Sox
21: 27660 Seattle_Mariners
22: 4340 Baltimore_Orioles
23: 4848143 New_York_Yankees
24: 30847 Toronto_Blue_Jays
25: 8579 Detroit_Tigers
26: 16883 Kansas_City_Royals
27: 30857 Texas_Rangers_(baseball)
28: 20050 Minnesota_Twins
29: 2456294 National_League_East
30: 2140 Atlanta_Braves
31: 20153 Montreal_Expos
32: 4848143 New_York_Yankees
33: 23741 Philadelphia_Phillies
34: 2456377 National_League_Central
35: 13894 Houston_Astros
36: 6671 Cincinnati_Reds

37: 24270 Pittsburgh_Pirates
38: 2456474 National_League_West
39: 28417 San_Diego_Padres
40: 6669 Colorado_Rockies
41: 28416 San_Francisco_Giants
42: 6669 Colorado_Rockies
43: 6671 Cincinnati_Reds
44: 2140 Atlanta_Braves
45: 24270 Pittsburgh_Pirates
46: 13894 Houston_Astros
47: 23741 Philadelphia_Phillies
48: 20153 Montreal_Expos
49: 28416 San_Francisco_Giants
1: 6654 Chicago_Cubs
 [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.7288182, final_local_scores=-0.81832755,
pem=-4.591283, similarity=4.6372952|
 RECALL MISS: 6654 Chicago_Cubs]
2: 1360083 Los_Angeles_Angels_ofAnaheim
 [winner: 719053 California_State_Route_58, pem_pos=0,
final_scores=-0.8082617, final_local_scores=-0.75627416,
pem=0.0, similarity=-18.479017|
 RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
3: 1360083 Los_Angeles_Angels_ofAnaheim
 [winner: 5407 California, pem_pos=0,
final_scores=-0.7855398, final_local_scores=-0.8125932,
pem=-0.08662026, similarity=-19.688635|
 RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
4: 6654 Chicago_Cubs
 [winner: 6886 Chicago, pem_pos=0,
final_scores=-0.6427981, final_local_scores=-0.69708496,
pem=-0.20658566, similarity=-15.589249|
 RECALL MISS: 6654 Chicago_Cubs]
5: 11278 Miami_Marlins
 [winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.78607696, final_local_scores=-0.8602794,
pem=-4.620281, similarity=3.5430675|
 RECALL MISS: 11278 Miami_Marlins]
6: 5945 Chicago_White_Sox
 [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.7495704, final_local_scores=-0.84099627,
pem=-4.591283, similarity=3.960816|
 RECALL MISS: 5945 Chicago_White_Sox]
7: 1360083 Los_Angeles_Angels_ofAnaheim
 [winner: 18213 Los_Angeles_Dodgers, pem_pos=15,
final_scores=-0.21123578, final_local_scores=-0.3459174,
pem=-4.7524934, similarity=19.611647|
 RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
8: 1360083 Los_Angeles_Angels_ofAnaheim
 [winner: 18110 Los_Angeles, pem_pos=0,
final_scores=-0.012282684, final_local_scores=0.011497408,
pem=-0.100652285, similarity=4.980181|

RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]

1: 38776 Major_League_Baseball
[pem_pos=0, final_scores=0.5869629,
final_local_scores=0.5392929, pem=-0.47792417,
similarity=22.782312]

2: 645042 New_York_City
[pem_pos=1, final_scores=0.3671915,
final_local_scores=0.43480438, pem=-1.7906687,
similarity=26.803078]

3: 38776 Major_League_Baseball
[pem_pos=0, final_scores=0.53944844,
final_local_scores=0.48739058, pem=0.0, similarity=18.63442]

4: 2004582 American_League_East
[pem_pos=0, final_scores=0.4788956,
final_local_scores=0.47675347, pem=0.0, similarity=18.316988]

5: 4848143 New_York_Yankees
[pem_pos=11, final_scores=0.52114314,
final_local_scores=0.4562791, pem=-5.0692368,
similarity=45.27329]

6: 4340 Baltimore_Orioles
[pem_pos=3, final_scores=1.2799306,
final_local_scores=1.2822444, pem=-4.046561,
similarity=64.36027]

7: 4339 Boston_Red_Sox
[pem_pos=5, final_scores=1.179256,
final_local_scores=1.1763291, pem=-4.418931,
similarity=63.224556]

8: 30847 Toronto_Blue_Jays
[pem_pos=6, final_scores=0.7083443,
final_local_scores=0.6671901, pem=-4.422637,
similarity=48.050983]

9: 8579 Detroit_Tigers
[pem_pos=4, final_scores=1.2647595,
final_local_scores=1.2659051, pem=-4.1401463,
similarity=64.38161]

10: 2456377 National_League_Central
[pem_pos=0, final_scores=0.57224244,
final_local_scores=0.5893182, pem=-0.98951703,
similarity=27.057293]

11: 6652 Cleveland_Indians
[pem_pos=12, final_scores=1.2774725,
final_local_scores=1.2792962, pem=-4.191953,
similarity=65.06296]

12: 20050 Minnesota_Twins
[pem_pos=9, final_scores=1.2772496,
final_local_scores=1.2819042, pem=-4.2906566,
similarity=65.67755]

13: 20152 Milwaukee_Brewers
[pem_pos=4, final_scores=0.48883975,
final_local_scores=0.4228049, pem=-3.40205,
similarity=35.20793]

14: 16883 Kansas_City_Royals

```
[pem_pos=5, final_scores=0.8360432,
final_local_scores=0.8024396, pem=-3.5915844,
similarity=47.567703]
15: 2456474 National_League_West
    [pem_pos=1, final_scores=0.5638904,
final_local_scores=0.56669587, pem=-1.2952092,
similarity=28.044596]
16: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.3169777,
final_local_scores=1.3266416, pem=-4.928425,
similarity=70.48089]
17: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.72941154,
final_local_scores=0.6913158, pem=-3.8567436,
similarity=45.693527]
18: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.318246,
final_local_scores=1.3247545, pem=-3.9188845,
similarity=64.93453]
19: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.2148044,
final_local_scores=1.2117603, pem=-3.9188845,
similarity=61.56256]
20: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=0.95936507,
final_local_scores=0.93613154, pem=-4.418931,
similarity=56.05658]
21: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.706711,
final_local_scores=0.666519, pem=-3.8567436,
similarity=44.95354]
22: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.1968946,
final_local_scores=1.1915402, pem=-4.046561,
similarity=61.653473]
23: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.31400067,
final_local_scores=0.23025614, pem=-5.0692368,
similarity=38.528313]
24: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.6718082,
final_local_scores=0.6272799, pem=-4.422637,
similarity=46.859985]
25: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.1919897,
final_local_scores=1.1864152, pem=-4.1401463,
similarity=62.00947]
26: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.8000166,
final_local_scores=0.7630861, pem=-3.5915844,
similarity=46.39332]
27: 30857 Texas_Rangers_(baseball)
```

```
[pem_pos=13, final_scores=1.2054598,
final_local_scores=1.2048252, pem=-4.928425,
similarity=66.84564]
28: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=0.9364479,
final_local_scores=0.9096301, pem=-4.2906566,
similarity=54.56815]
29: 2456294 National_League_East
    [pem_pos=0, final_scores=0.43432057,
final_local_scores=0.42720723, pem=0.0, similarity=16.838428]
30: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=1.1620841,
final_local_scores=1.153328, pem=-4.136474,
similarity=61.00211]
31: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.83475405,
final_local_scores=0.80227375, pem=-4.5098605,
similarity=52.556484]
32: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.55154324,
final_local_scores=0.48948663, pem=-5.0692368,
similarity=46.264267]
33: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.2769336,
final_local_scores=1.2801452, pem=-4.5189333,
similarity=66.86646]
34: 2456377 National_League_Central
    [pem_pos=0, final_scores=0.56933373,
final_local_scores=0.58614093, pem=-0.98951703,
similarity=26.962475]
35: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.3091757,
final_local_scores=1.3157732, pem=-4.650927,
similarity=68.64748]
36: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.0863414,
final_local_scores=1.0718312, pem=-3.771247,
similarity=56.58392]
37: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.303969,
final_local_scores=1.3079083, pem=-4.03702,
similarity=65.07425]
38: 2456474 National_League_West
    [pem_pos=1, final_scores=0.5752174,
final_local_scores=0.57906896, pem=-1.2952092,
similarity=28.413834]
39: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.91718715,
final_local_scores=0.89216906, pem=-3.9307637,
similarity=52.089924]
40: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.1335131,
```

```
final_local_scores=1.1311328, pem=-4.482383,
similarity=62.220867]
41: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.96474725,
final_local_scores=0.9381588, pem=-4.409061,
similarity=56.0634]
42: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.0379735,
final_local_scores=1.0267704, pem=-4.482383,
similarity=59.106483]
43: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.2381622,
final_local_scores=1.2376723, pem=-3.771247,
similarity=61.53295]
44: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=1.0006267,
final_local_scores=0.9769601, pem=-4.136474,
similarity=55.73894]
45: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.2899419,
final_local_scores=1.2925858, pem=-4.03702, similarity=64.617]
46: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.2093697,
final_local_scores=1.2067504, pem=-4.650927,
similarity=65.39402]
47: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.3255701,
final_local_scores=1.3332732, pem=-4.5189333,
similarity=68.451904]
48: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.61077124,
final_local_scores=0.5576066, pem=-4.5098605,
similarity=45.255127]
49: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.82299405,
final_local_scores=0.783315, pem=-4.409061,
similarity=51.442554]
```

Sample: b'52_BASEBALL'

```
BASEBALL - [1 1 MAJOR LEAGUE 1] RESULTS WEDNESDAY . [2 2 NEW
YORK 2] 1996-08-22 Results of [3 3 Major League 3] Baseball
games played on Wednesday ( home team in CAPS ) : [4 4 Americ
an League 4] 1 California 7 [5 5 NEW YORK 5] 1 [6 6 DETROIT 6]
7 2 Chicago 4 [7 7 Milwaukee 7] 10 [8 8 MINNESOTA 8] 7 [9 9 BO
STON 9] 6 [10 10 Oakland 10] 4 [11 11 BALTIMORE 11] 10 [12 12
Seattle 12] 5 [13 13 Texas 13] 10 [14 14 CLEVELAND 14] 8 ( in
10 ) [15 15 Toronto 15] 6 [16 16 KANSAS CITY 16] 2 [17 17 Nat
ional League 17] 3 CHICAGO 8 4 Florida 3 [18 18 SAN FRANCISCO
18] 12 [19 19 New York 19] 11 [20 20 ATLANTA 20] 4 [21 21 Cin
cinnati 21] 3 [22 22 Pittsburgh 22] 5 [23 23 HOUSTON 23] 2 [24
24 COLORADO 24] 10 St Louis 2 [25 25 Philadelphia 25] 6 5 LOS
ANGELES 0 [26 26 SAN DIEGO 26] 7 [27 27 Montreal 27] 2
```

```
1: 38776 Major_League_Baseball
2: 645042 New_York_City
3: 38776 Major_League_Baseball
4: 64777 American_League
5: 4848143 New_York_Yankees
6: 8579 Detroit_Tigers
7: 20152 Milwaukee_Brewers
8: 20050 Minnesota_Twins
9: 4339 Boston_Red_Sox
10: 22523 Oakland_Athletics
11: 4340 Baltimore_Orioles
12: 27660 Seattle_Mariners
13: 30857 Texas_Rangers_(baseball)
14: 6652 Cleveland_Indians
15: 30847 Toronto_Blue_Jays
16: 16883 Kansas_City_Royals
17: 64290 National_League
18: 28416 San_Francisco_Giants
19: 4848143 New_York_Yankees
20: 2140 Atlanta_Braves
21: 6671 Cincinnati_Reds
22: 24270 Pittsburgh_Pirates
23: 13894 Houston_Astros
24: 6669 Colorado_Rockies
25: 23741 Philadelphia_Phillies
26: 28417 San_Diego_Padres
27: 20153 Montreal_Expos
1: 1360083 Los_Angeles_Angels_of_Anheim
    [winner: 7590031 California_Golden_Bears_football,
pem_pos=1, final_scores=-0.67227614,
final_local_scores=-0.73627037, pem=-4.4771376,
similarity=6.465308|
     RECALL MISS: 1360083 Los_Angeles_Angels_of_Anheim]
2: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.5761967, final_local_scores=-0.64963627,
pem=-4.591283, similarity=9.67138|
     RECALL MISS: 5945 Chicago_White_Sox]
3: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.7754649, final_local_scores=-0.86730653,
pem=-4.591283, similarity=3.1756659|
     RECALL MISS: 5945 Chicago_White_Sox]
4: 11278 Miami_Marlins
    [winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.7923735, final_local_scores=-0.8642763,
pem=-4.620281, similarity=3.4237921|
     RECALL MISS: 11278 Miami_Marlins]
5: 1360083 Los_Angeles_Angels_of_Anheim
    [winner: 18213 Los_Angeles_Dodgers, pem_pos=15,
final_scores=-0.23051299, final_local_scores=-0.36698723,
pem=-4.7524934, similarity=18.98288|
```

RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]

1: 38776 Major_League_Baseball
[pem_pos=0, final_scores=0.5457727,
final_local_scores=0.4915172, pem=-0.47792417,
similarity=21.356592]
2: 645042 New_York_City
[pem_pos=1, final_scores=0.2575376,
final_local_scores=0.3225717, pem=-1.7906687,
similarity=23.45383]
3: 38776 Major_League_Baseball
[pem_pos=0, final_scores=0.5550605,
final_local_scores=0.50166273, pem=-0.47792417,
similarity=21.659355]
4: 64777 American_League
[pem_pos=0, final_scores=0.5559597,
final_local_scores=0.5254387, pem=0.0, similarity=19.769852]
5: 4848143 New_York_Yankees
[pem_pos=11, final_scores=0.43424928,
final_local_scores=0.3614064, pem=-5.0692368,
similarity=42.442097]
6: 8579 Detroit_Tigers
[pem_pos=4, final_scores=1.2058725,
final_local_scores=1.200645, pem=-4.1401463,
similarity=62.434113]
7: 20152 Milwaukee_Brewers
[pem_pos=4, final_scores=0.37866813,
final_local_scores=0.3021589, pem=-3.40205,
similarity=31.607615]
8: 20050 Minnesota_Twins
[pem_pos=9, final_scores=0.9962955,
final_local_scores=0.9737242, pem=-4.2906566,
similarity=56.480843]
9: 4339 Boston_Red_Sox
[pem_pos=5, final_scores=1.0609534,
final_local_scores=1.047003, pem=-4.418931,
similarity=59.365204]
10: 22523 Oakland_Athletics
[pem_pos=7, final_scores=0.9486106,
final_local_scores=0.92051, pem=-3.9188845,
similarity=52.87107]
11: 4340 Baltimore_Orioles
[pem_pos=3, final_scores=1.2523005,
final_local_scores=1.2510958, pem=-4.046561,
similarity=63.430733]
12: 27660 Seattle_Mariners
[pem_pos=4, final_scores=0.49390906,
final_local_scores=0.43430477, pem=-3.8567436,
similarity=38.023804]
13: 30857 Texas_Rangers_(baseball)
[pem_pos=13, final_scores=0.9047983,
final_local_scores=0.8764519, pem=-4.928425,
similarity=57.046326]

```
14: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.2391233,
final_local_scores=1.2360945, pem=-4.191953,
similarity=63.773735]
15: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.9046585,
final_local_scores=0.8823641, pem=-4.422637,
similarity=54.472202]
16: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.7568616,
final_local_scores=0.71540487, pem=-3.5915844,
similarity=44.970417]
17: 64290 National_League
    [pem_pos=0, final_scores=0.58202076,
final_local_scores=0.5487162, pem=-0.07431644,
similarity=20.868643]
18: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.85749954,
final_local_scores=0.82098323, pem=-4.409061,
similarity=52.56665]
19: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.5828503,
final_local_scores=0.52373046, pem=-5.0692368,
similarity=47.28617]
20: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=0.92698616,
final_local_scores=0.8966511, pem=-4.136474,
similarity=53.34236]
21: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=0.9944139,
final_local_scores=0.9707388, pem=-3.771247,
similarity=53.567123]
22: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=0.9873753,
final_local_scores=0.9612401, pem=-4.03702,
similarity=54.728977]
23: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.218589,
final_local_scores=1.2173839, pem=-4.650927,
similarity=65.71134]
24: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.0760224,
final_local_scores=1.0696101, pem=-4.482383,
similarity=60.384903]
25: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.0339861,
final_local_scores=1.0142009, pem=-4.5189333,
similarity=58.930153]
26: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.75750047,
final_local_scores=0.7178055, pem=-3.9307637,
similarity=46.886566]
```

```
27: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.57715255,
final_local_scores=0.52166426, pem=-4.5098605,
similarity=44.182537]
```

Sample: b'53_BASEBALL'

BASEBALL - [1 1 GREER 1] HOMER IN 10TH LIFTS [2 2 TEXAS 2]
PAST [3 3 INDIANS 3] . [4 4 CLEVELAND 4] 1996-08-22 [5 5 Rusty
Greer 5] 's two-run homer in the top of the 10th inning
rallied the [6 6 Texas Rangers 6] to a 10-8 victory over the [7 7 Cleveland Indians 7] Wednesday in the rubber game of a
three-game series between division leaders . With one out , [8
8 Greer 8] hit a 1-1 pitch from [9 9 Julian Tavarez 9]
(4-7) over the right-field fence for his 15th home run . "
It was an off-speed pitch and I just tried to get a good
swing on it and put it in play , " [10 10 Greer 10] said . "
This was a big game . The crowd was behind him and it was
intense . " The shot brought home [11 11 Ivan Rodriguez 11] ,
who had his second double of the game , giving him 42 this
season , 41 as a catcher . He joined [12 12 Mickey Cochrane 12]
, [13 13 Johnny Bench 13] and [14 14 Terry Kennedy 14] as
the only catchers with 40 doubles in a season . The [15 15 Ra
ngers 15] have won 10 of their last 12 games and six of nine
meetings against the [16 16 Indians 16] this season . The
American League Western leaders have won eight of 15 games at
[17 17 Jacobs Field 17] , joining the [18 18 Yankees 18] as
the only teams with a winning record at the A.L. Central
leaders ' home . [19 19 Cleveland 19] lost for just the
second time in six games . The [20 20 Indians 20] sent the
game into extra innings in the ninth on [21 21 Kenny Lofton 21]
's two-run single . [22 22 Ed Vosberg 22] (1-0) blew his
first save opportunity but got the win , allowing three hits
with two walks and three strikeouts in 1 2/3 scoreless
innings . [23 23 Dean Palmer 23] hit his 30th homer for the [2
4 24 Rangers 24] . In [25 25 Baltimore 25] , [26 26 Cal Ripken
26] had four hits and snapped a fifth-inning tie with a solo
homer and [27 27 Bobby Bonilla 27] added a three-run shot in
the seventh to power the surging [28 28 Orioles 28] to a 10-5
victory over the [29 29 Seattle Mariners 29] . The [30 30 Mar
iners 30] scored four runs in the top of the fifth to tie the
game 5-5 but [31 31 Ripken 31] led off the bottom of the
inning with his 21st homer off starter [32 32 Sterling Hitchc
ock 32] (12-6) . [33 33 Bonilla 33] 's blast was the first
time [34 34 Randy Johnson 34] , last season 's [35 35 Cy Young
35] winner , allowed a run in five relief appearances since
coming off the disabled list on August 6 . [36 36 Bonilla 36]
has 21 RBI and 15 runs in his last 20 games . [37 37 Baltimore
37] has won seven of nine and 16 of its last 22 and cut the [3
8 38 Yankees 38] ' lead in the 1 A.L. East to five games . [39
39 Scott Erickson 39] (8-10) laboured to his third straight
win . [40 40 Alex Rodriguez 40] had two homers and four RBI
for the [41 41 Mariners 41] , who have dropped three in a row

and 11 of 15 . He became the fifth shortstop in major-league history to hit 30 homers in a season and the first since [42 42 Ripken 42] hit 34 in 1991 . [43 43 Chris Hoiles 43] hit his 22nd homer for [44 44 Baltimore 44] . In [45 45 New York 45] , [46 46 Jason Dickson 46] scattered 10 hits over 6 1/3 innings in his major-league debut and [47 47 Chili Davis 47] belted a homer from each side of the plate as the [48 48 California Angels 48] defeated the [49 49 Yankees 49] 7-1 . [50 50 Dickson 50] allowed a homer to [51 51 Derek Jeter 51] on his first major-league pitch but settled down . He was the 27th pitcher used by the [52 52 Angels 52] this season , tying a major-league record . [53 53 Jimmy Key 53] (9-10) took the loss as the [54 54 Yankees 54] lost their ninth in 14 games . They stranded 11 baserunners . 2 California played without interim manager [55 55 John McNamara 55] , who was admitted to a [56 56 New York 56] hospital with a blood clot in his right calf . In [57 57 Boston 57] , [58 58 Mike Stanley 58] 's bases-loaded two-run single snapped an eighth-inning tie and gave the [59 59 Red Sox 59] their third straight win , 6-4 over the [60 60 Oakland Athletics 60] . Stanley owns a .367 career batting average with the bases loaded (33-for-90) . [61 61 Boston 61] 's [62 62 Mo Vaughn 62] went 3-for-3 with a walk , stole home for one of his three runs scored and collected his 116th RBI . [63 63 Scott Brosius 63] homered and drove in two runs for the [64 64 Athletics 64] , who have lost seven of their last nine games . In [65 65 Detroit 65] , [66 66 Brad Ausmus 66] 's three-run homer capped a four-run eighth and lifted the [67 67 Tigers 67] to a 7-4 victory over the reeling [68 68 Chicago White Sox 68] . The [69 69 Tigers 69] have won consecutive games after dropping eight in a row , but have won nine of their last 12 at home . The [70 70 White Sox 70] have lost six of their last eight games . In [71 71 Kansas City 71] , [72 72 Juan Guzman 72] tossed a complete-game six-hitter to win for the first time in over a month and lower his league-best ERA as the [73 73 Toronto Blue Jays 73] won their fourth straight , 6-2 over the [74 74 Royals 74] . Guzman (10-8) won for the first time since July 16 , a span of six starts . He allowed two runs -- one earned -- and lowered his ERA to 2.99 . At [75 75 Minnesota 75] , [76 76 John Jaha 76] 's three-run homer , his 26th , capped a five-run eighth inning that rallied the [77 77 Milwaukee Brewers 77] to a 10-7 victory over the [78 78 Twins 78] . Jaha added an RBI single in the ninth and had four RBI . [79 79 Jose Valentin 79] hit his 21st homer for [80 80 Milwaukee 80] .

- 1: 2070208 Rusty_Greer
- 2: 30857 Texas_Rangers_(baseball)
- 3: 6652 Cleveland_Indians
- 4: 5951 Cleveland
- 5: 2070208 Rusty_Greer
- 6: 30857 Texas_Rangers_(baseball)
- 7: 6652 Cleveland_Indians

8: 2070208 Rusty_Greer
9: 2407111 Julián_Tavárez
10: 2070208 Rusty_Greer
11: 552153 Iván_Rodríguez
12: 161226 Mickey_Cochrane
13: 16581 Johnny_Bench
14: 1261670 Terry_Kennedy
15: 30857 Texas_Rangers_(baseball)
16: 6652 Cleveland_Indians
17: 329401 Progressive_Field
18: 4848143 New_York_Yankees
19: 6652 Cleveland_Indians
20: 6652 Cleveland_Indians
21: 1264752 Kenny_Lofton
22: 2733958 Ed_Vosberg
23: 2644877 Dean_Palmer
24: 30857 Texas_Rangers_(baseball)
25: 4340 Baltimore_Orioles
26: 153856 Cal_Ripken,_Jr.
27: 905899 Bobby_Bonilla
28: 4340 Baltimore_Orioles
29: 27660 Seattle_Mariners
30: 27660 Seattle_Mariners
31: 153856 Cal_Ripken,_Jr.
32: 6333369 Sterling_Hitchcock
33: 905899 Bobby_Bonilla
34: 386384 Randy_Johnson
35: 6725 Cy_Young_Award
36: 905899 Bobby_Bonilla
37: 4340 Baltimore_Orioles
38: 4848143 New_York_Yankees
39: 1915655 Scott_Erickson
40: 212472 Alex_Rodriguez
41: 27660 Seattle_Mariners
42: 153856 Cal_Ripken,_Jr.
43: 2765457 Chris_Hoiles
44: 4340 Baltimore_Orioles
45: 645042 New_York_City
46: 931528 Jason_Dickson
47: 1547794 Chili_Davis
48: 1360083 Los_Angeles_Angels_of_Anheim
49: 4848143 New_York_Yankees
50: 931528 Jason_Dickson
51: 218947 Derek_Jeter
52: 1360083 Los_Angeles_Angels_of_Anheim
53: 385773 Jimmy_Key
54: 4848143 New_York_Yankees
55: 2226049 John_McNamara_(baseball)
56: 645042 New_York_City
57: 24437894 Boston
58: 2400192 Mike_Stanley
59: 4339 Boston_Red_Sox

```
60: 22523 Oakland_Athletics
61: 4339 Boston_Red_Sox
62: 395171 Mo_Vaughn
63: 877552 Scott_Brosius
64: 22523 Oakland_Athletics
65: 8687 Detroit
66: 1838414 Brad_Ausmus
67: 8579 Detroit_Tigers
68: 5945 Chicago_White_Sox
69: 8579 Detroit_Tigers
70: 5945 Chicago_White_Sox
71: 17454 Kansas_City,_Missouri
72: 3139393 Juan_Guzmán_(baseball)
73: 30847 Toronto_Blue_Jays
74: 16883 Kansas_City_Royals
75: 19590 Minnesota
76: 2787747 John_Jaha
77: 20152 Milwaukee_Brewers
78: 20050 Minnesota_Twins
79: 2406308 José_Valentín
80: 20152 Milwaukee_Brewers
1: 2004582 American_League_East
    [winner: 64777 American_League, pem_pos=0,
final_scores=-1.2185849, final_local_scores=-1.4143817,
pem=0.0, similarity=-38.11826]
    RECALL MISS: 2004582 American_League_East]
    [winner: 31011522
National_Register_of_Historic_Places_listings_in_east_Davenport,_Iowa, pem_pos=17, final_scores=-1.2511718,
final_local_scores=-1.1695251, pem=-4.151308,
similarity=-8.235794]
    gt: 2004582 American_League_East, pem_pos=14,
final_scores=-1.4508874, final_local_scores=-1.6308887,
pem=-4.0594826, similarity=-22.503164 ]
2: 1360083 Los_Angeles_Angels_of_Anheim
    [winner: 7590031 California_Golden_Bears_football,
pem_pos=1, final_scores=-0.6096951,
final_local_scores=-0.66409343, pem=-4.4771376,
similarity=8.619211]
    RECALL MISS: 1360083 Los_Angeles_Angels_of_Anheim]
1: 2070208 Rusty_Greer
    [pem_pos=0, final_scores=0.5296047,
final_local_scores=0.57865626, pem=0.0, similarity=21.35797]
2: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.081192,
final_local_scores=1.073045, pem=-4.928425,
similarity=62.91306]
3: 6652 Cleveland_Indians
    [pem_pos=3, final_scores=0.695869,
final_local_scores=0.64935416, pem=-2.3830547,
similarity=36.42716]
4: 5951 Cleveland
```

```
[pem_pos=0, final_scores=0.4054628,
final_local_scores=0.45483017, pem=-0.57246673,
similarity=20.775915]
5: 2070208 Rusty_Greer
    [pem_pos=0, final_scores=0.564801,
final_local_scores=0.6171029, pem=0.0, similarity=22.505295]
6: 30857 Texas_Rangers_(baseball)
    [pem_pos=0, final_scores=0.6070873,
final_local_scores=0.55515766, pem=-0.15069729,
similarity=21.47624]
7: 6652 Cleveland_Indians
    [pem_pos=0, final_scores=0.5057594,
final_local_scores=0.4416883, pem=-0.08813761,
similarity=17.74988]
8: 2070208 Rusty_Greer
    [pem_pos=0, final_scores=0.6517709,
final_local_scores=0.7121043, pem=0.0, similarity=25.340326]
9: 2407111 Julián_Tavárez
    [pem_pos=0, final_scores=0.65113306,
final_local_scores=0.68988997, pem=0.0, similarity=24.677406]
10: 2070208 Rusty_Greer
    [pem_pos=0, final_scores=0.6824077,
final_local_scores=0.74557036, pem=0.0, similarity=26.33902]
11: 552153 Iván_Rodríguez
    [pem_pos=0, final_scores=0.66296536,
final_local_scores=0.6707867, pem=0.0, similarity=24.107325]
12: 161226 Mickey_Cochrane
    [pem_pos=0, final_scores=0.44738114,
final_local_scores=0.45736247, pem=0.0, similarity=17.738321]
13: 16581 Johnny_Bench
    [pem_pos=0, final_scores=0.54261297,
final_local_scores=0.53792447, pem=0.0, similarity=20.142452]
14: 1261670 Terry_Kennedy
    [pem_pos=0, final_scores=0.56291413,
final_local_scores=0.5850341, pem=0.0, similarity=21.548296]
15: 30857 Texas_Rangers_(baseball)
    [pem_pos=1, final_scores=0.53878695,
final_local_scores=0.48054987, pem=-2.6930556,
similarity=33.075535]
16: 6652 Cleveland_Indians
    [pem_pos=3, final_scores=0.6244289,
final_local_scores=0.57131666, pem=-2.3830547,
similarity=34.098366]
17: 329401 Progressive_Field
    [pem_pos=0, final_scores=0.54743785,
final_local_scores=0.58121, pem=0.0, similarity=21.43418]
18: 4848143 New_York_Yankees
    [pem_pos=0, final_scores=0.54876417,
final_local_scores=0.48903805, pem=-0.4294726,
similarity=21.019123]
19: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=0.81791824,
```

```
final_local_scores=0.7826744, pem=-4.191953,
similarity=50.24277]
20: 6652 Cleveland_Indians
    [pem_pos=3, final_scores=0.50246054,
final_local_scores=0.43808478, pem=-2.3830547,
similarity=30.122461]
21: 1264752 Kenny_Lofton
    [pem_pos=0, final_scores=0.77669984,
final_local_scores=0.78957856, pem=0.0, similarity=27.652311]
22: 2733958 Ed_Vosberg
    [pem_pos=0, final_scores=0.516001,
final_local_scores=0.5805292, pem=0.0, similarity=21.413862]
23: 2644877 Dean_Palmer
    [pem_pos=0, final_scores=0.5810644,
final_local_scores=0.62947094, pem=0.0, similarity=22.874382]
24: 30857 Texas_Rangers_(baseball)
    [pem_pos=1, final_scores=0.5863313,
final_local_scores=0.5324848, pem=-2.6930556,
similarity=34.625378]
25: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=0.5695782,
final_local_scores=0.5132093, pem=-4.046561,
similarity=41.41073]
26: 153856 Cal_Ripken,_Jr.
    [pem_pos=0, final_scores=0.44087762,
final_local_scores=0.42840564, pem=0.0, similarity=16.874191]
27: 905899 Bobby_Bonilla
    [pem_pos=0, final_scores=0.7044474,
final_local_scores=0.7299219, pem=0.0, similarity=25.872038]
28: 4340 Baltimore_Orioles
    [pem_pos=0, final_scores=0.49537122,
final_local_scores=0.43214947, pem=-0.18171716,
similarity=17.974121]
29: 27660 Seattle_Mariners
    [pem_pos=0, final_scores=0.4990371,
final_local_scores=0.44163138, pem=-0.008242749,
similarity=17.3137]
30: 27660 Seattle_Mariners
    [pem_pos=0, final_scores=0.4591099,
final_local_scores=0.398017, pem=-0.1261296,
similarity=16.653246]
31: 153856 Cal_Ripken,_Jr.
    [pem_pos=0, final_scores=0.39353353,
final_local_scores=0.37668937, pem=0.0, similarity=15.330874]
32: 6333369 Sterling_Hitchcock
    [pem_pos=0, final_scores=0.78739756,
final_local_scores=0.8313903, pem=0.0, similarity=28.900059]
33: 905899 Bobby_Bonilla
    [pem_pos=0, final_scores=0.73229104,
final_local_scores=0.76269084, pem=0.0, similarity=26.84993]
34: 386384 Randy_Johnson
    [pem_pos=0, final_scores=0.6262285,
```

```
final_local_scores=0.61475116, pem=0.0, similarity=22.435114]
35: 6725 Cy_Young_Award
    [pem_pos=1, final_scores=0.2771039,
final_local_scores=0.23428398, pem=-0.80355984,
similarity=15.4510975]
36: 905899 Bobby_Bonilla
    [pem_pos=0, final_scores=0.6147712,
final_local_scores=0.63431823, pem=0.0, similarity=23.019033]
37: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.0212975,
final_local_scores=1.006644, pem=-4.046561,
similarity=56.135803]
38: 4848143 New_York_Yankees
    [pem_pos=0, final_scores=0.55504435,
final_local_scores=0.4958982, pem=-0.4294726,
similarity=21.223843]
39: 1915655 Scott_Erickson
    [pem_pos=0, final_scores=0.7055206,
final_local_scores=0.72380984, pem=0.0, similarity=25.689642]
40: 212472 Alex_Rodriguez
    [pem_pos=0, final_scores=0.6105308,
final_local_scores=0.6030047, pem=0.0, similarity=22.084576]
41: 27660 Seattle_Mariners
    [pem_pos=0, final_scores=0.5097671,
final_local_scores=0.45335227, pem=-0.1261296,
similarity=18.304562]
42: 153856 Cal_Ripken,_Jr.
    [pem_pos=0, final_scores=0.41796738,
final_local_scores=0.40337968, pem=0.0, similarity=16.127367]
43: 2765457 Chris_Hoiles
    [pem_pos=0, final_scores=0.57878983,
final_local_scores=0.59954, pem=0.0, similarity=21.981182]
44: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=0.9195592,
final_local_scores=0.8955104, pem=-4.046561,
similarity=52.819355]
45: 645042 New_York_City
    [pem_pos=1, final_scores=0.4389608,
final_local_scores=0.51727927, pem=-1.7906687,
similarity=29.264292]
46: 931528 Jason_Dickson
    [pem_pos=0, final_scores=0.67800367,
final_local_scores=0.7888177, pem=0.0, similarity=27.629606]
47: 1547794 Chili_Davis
    [pem_pos=0, final_scores=0.6441615,
final_local_scores=0.6619341, pem=0.0, similarity=23.843145]
48: 1360083 Los_Angeles_Angels_of_Anheim
    [pem_pos=0, final_scores=0.4935841,
final_local_scores=0.4317974, pem=-0.016987266,
similarity=17.067787]
49: 4848143 New_York_Yankees
    [pem_pos=0, final_scores=0.47825116,
```

```
final_local_scores=0.41201323, pem=-0.4294726,
similarity=18.720547]
50: 931528 Jason_Dickson
    [pem_pos=0, final_scores=0.58182037,
final_local_scores=0.68375206, pem=0.0, similarity=24.494238]
51: 218947 Derek_Jeter
    [pem_pos=0, final_scores=0.55194336,
final_local_scores=0.5422288, pem=0.0, similarity=20.270903]
52: 1360083 Los_Angeles_Angels_of_Anheim
    [pem_pos=2, final_scores=0.5365705,
final_local_scores=0.47875345, pem=-1.7538311,
similarity=27.914276]
53: 385773 Jimmy_Key
    [pem_pos=0, final_scores=0.6604132,
final_local_scores=0.68550557, pem=0.0, similarity=24.546566]
54: 4848143 New_York_Yankees
    [pem_pos=0, final_scores=0.52196944,
final_local_scores=0.45976883, pem=-0.4294726,
similarity=20.145672]
55: 2226049 John_McNamara_(baseball)
    [pem_pos=0, final_scores=0.73837274,
final_local_scores=0.765639, pem=-0.0888523,
similarity=27.4211]
56: 645042 New_York_City
    [pem_pos=1, final_scores=0.37882233,
final_local_scores=0.45158714, pem=-1.7906687,
similarity=27.30391]
57: 24437894 Boston
    [pem_pos=0, final_scores=0.52087814,
final_local_scores=0.6037457, pem=-0.23164093,
similarity=23.366388]
58: 2400192 Mike_Stanley
    [pem_pos=0, final_scores=0.26213706,
final_local_scores=0.2855783, pem=0.0, similarity=12.611939]
59: 4339 Boston_Red_Sox
    [pem_pos=0, final_scores=0.46968657,
final_local_scores=0.4053973, pem=-0.07716359,
similarity=16.607204]
60: 22523 Oakland_Athletics
    [pem_pos=0, final_scores=0.5965968,
final_local_scores=0.5450756, pem=-0.06842491,
similarity=20.72796]
61: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=0.94056946,
final_local_scores=0.9197654, pem=-4.418931,
similarity=55.56818]
62: 395171 Mo_Vaughn
    [pem_pos=0, final_scores=0.8411104,
final_local_scores=0.8722382, pem=0.0, similarity=30.11904]
63: 877552 Scott_Brosius
    [pem_pos=0, final_scores=0.77884763,
final_local_scores=0.81675965, pem=0.0, similarity=28.46345]
```

```
64: 22523 Oakland_Athletics
    [pem_pos=2, final_scores=0.6628826,
final_local_scores=0.6174828, pem=-2.8318815,
similarity=37.916843]
65: 8687 Detroit
    [pem_pos=0, final_scores=0.516731,
final_local_scores=0.5749302, pem=-0.21700948,
similarity=22.426908]
66: 1838414 Brad_Ausmus
    [pem_pos=0, final_scores=0.8684465,
final_local_scores=0.8853542, pem=0.0, similarity=30.510448]
67: 8579 Detroit_Tigers
    [pem_pos=0, final_scores=0.54533285,
final_local_scores=0.48801178, pem=-1.1469651,
similarity=24.890333]
68: 5945 Chicago_White_Sox
    [pem_pos=0, final_scores=0.49219626,
final_local_scores=0.4273643, pem=-0.07886935,
similarity=17.272018]
69: 8579 Detroit_Tigers
    [pem_pos=0, final_scores=0.6700781,
final_local_scores=0.62427706, pem=-1.1469651,
similarity=28.95676]
70: 5945 Chicago_White_Sox
    [pem_pos=0, final_scores=0.595549,
final_local_scores=0.5402614, pem=-0.08007053,
similarity=20.647627]
71: 17454 Kansas_City,_Missouri
    [pem_pos=0, final_scores=0.35531992,
final_local_scores=0.38659412, pem=-0.8046767,
similarity=20.002409]
72: 3139393 Juan_Guzmán_(baseball)
    [pem_pos=0, final_scores=0.3861515,
final_local_scores=0.40522254, pem=0.0, similarity=16.182362]
73: 30847 Toronto_Blue_Jays
    [pem_pos=0, final_scores=0.48622847,
final_local_scores=0.42886448, pem=-0.018220792,
similarity=16.98697]
74: 16883 Kansas_City_Royals
    [pem_pos=0, final_scores=0.53628254,
final_local_scores=0.47895575, pem=-0.31042016,
similarity=20.070822]
75: 19590 Minnesota
    [pem_pos=0, final_scores=0.493802,
final_local_scores=0.59269077, pem=-0.21388707,
similarity=22.939938]
76: 2787747 John_Jaha
    [pem_pos=0, final_scores=0.43879622,
final_local_scores=0.4995193, pem=0.0, similarity=18.996365]
77: 20152 Milwaukee_Brewers
    [pem_pos=0, final_scores=0.45243442,
final_local_scores=0.3885119, pem=-0.12176443,
```

```
similarity=16.345856]
78: 20050 Minnesota_Twins
    [pem_pos=1, final_scores=0.4414873,
final_local_scores=0.37776208, pem=-1.3860784,
similarity=22.900597]
79: 2406308 José_Valentín
    [pem_pos=0, final_scores=0.42695767,
final_local_scores=0.43359345, pem=0.0, similarity=17.029007]
80: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.38492042,
final_local_scores=0.31476313, pem=-3.40205,
similarity=31.98375]
```

Sample: b'54_SOCCER'

SOCCER - [1 1 COCU 1] DOUBLE EARNS [2 2 PSV 2] 4-1 WIN . [3 3 AMSTERDAM 3] 1996-08-22 [4 4 Philip Cocu 4] scored twice in the second half to spur [5 5 PSV Eindhoven 5] to a 4-1 away win over [6 6 NEC Nijmegen 6] in the [7 7 Dutch 7] first division on Thursday . He scored from close range in the 54th minute and from a bicycle kick 13 minutes later . [8 8 Arthur Numan 8] and [9 9 Luc Nilis 9] , [10 10 Dutch 10] top scorer last season , were [11 11 PSV 11] 's other marksmen . [12 12 Ajax Amsterdam 12] opened their title defence with a 1-0 win over [13 13 NAC Breda 13] on Wednesday .

```
1: 1295837 Phillip_Cocu
2: 79820 PSV_Eindhoven
3: 844 Amsterdam
4: 1295837 Phillip_Cocu
5: 79820 PSV_Eindhoven
6: 834154 N.E.C._(football_club)
7: 21148 Netherlands
8: 1218987 Arthur_Numan
9: 2945257 Luc_Nilis
10: 21148 Netherlands
11: 79820 PSV_Eindhoven
12: 2273 AFC_Ajax
13: 505173 NAC_Breda
1: 1295837 Phillip_Cocu
    [pem_pos=0, final_scores=0.8919466,
final_local_scores=0.91940683, pem=0.0, similarity=31.526646]
2: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.6599092,
final_local_scores=0.62103415, pem=0.0, similarity=22.62261]
3: 844 Amsterdam
    [pem_pos=0, final_scores=0.60465497,
final_local_scores=0.6517309, pem=-0.0642803,
similarity=23.88823]
4: 1295837 Phillip_Cocu
    [pem_pos=0, final_scores=0.88956326,
final_local_scores=0.9051848, pem=0.0, similarity=31.102234]
5: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.5278405,
```

```
final_local_scores=0.4767691, pem=0.0, similarity=18.317453]
6: 834154 N.E.C._(football_club)
    [pem_pos=0, final_scores=0.54494697,
final_local_scores=0.520771, pem=0.0, similarity=19.63056]
7: 21148 Netherlands
    [pem_pos=0, final_scores=0.4962514,
final_local_scores=0.5231859, pem=-0.7943452,
similarity=24.022396]
8: 1218987 Arthur_Numan
    [pem_pos=0, final_scores=0.78757113,
final_local_scores=0.8295136, pem=0.0, similarity=28.844053]
9: 2945257 Luc_Nilis
    [pem_pos=0, final_scores=0.8456231,
final_local_scores=0.93737644, pem=0.0, similarity=32.062897]
10: 21148 Netherlands
    [pem_pos=0, final_scores=0.38964254,
final_local_scores=0.4067319, pem=-0.7943452,
similarity=20.547176]
11: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.56772226,
final_local_scores=0.5203338, pem=0.0, similarity=19.617512]
12: 2273 AFC_Ajax
    [pem_pos=0, final_scores=0.60183847,
final_local_scores=0.5610423, pem=0.0, similarity=20.832335]
13: 505173 NAC_Breda
    [pem_pos=0, final_scores=0.5624108,
final_local_scores=0.54738927, pem=0.0, similarity=20.4249]
```

Sample: b'55_SOCCER'

```
SOCCER - [1 1 DUTCH 1] FIRST DIVISION SUMMARY . [2 2 AMSTERDAM 2] 1996-08-22 Summary of Thursday 's only [3 3 Dutch 3]
first division match : [4 4 NEC Nijmegen 4] 1 ( Van Eykeren
15th ) [5 5 PSV Eindhoven 5] 4 ( Numan 11th , [6 6 Nilis 6]
42nd , [7 7 Cocu 7] 54th , 67th ) . Halftime 1-2 . Attendance
8,000
1: 21148 Netherlands
2: 844 Amsterdam
3: 21148 Netherlands
4: 834154 N.E.C._(football_club)
5: 79820 PSV_Eindhoven
6: 2945257 Luc_Nilis
7: 1295837 Phillip_Cocu
1: 21148 Netherlands
    [pem_pos=0, final_scores=0.67406553,
final_local_scores=0.679447, pem=-0.7943452,
similarity=28.68554]
2: 844 Amsterdam
    [pem_pos=0, final_scores=0.6476385,
final_local_scores=0.66853446, pem=-0.0642803,
similarity=24.38968]
3: 21148 Netherlands
    [pem_pos=0, final_scores=0.545124,
```

```
final_local_scores=0.538598, pem=-0.7943452,
similarity=24.482325]
4: 834154 N.E.C._(football_club)
    [pem_pos=0, final_scores=0.6014648,
final_local_scores=0.614387, pem=0.0, similarity=22.424246]
5: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.47731876,
final_local_scores=0.45942205, pem=0.0, similarity=17.799784]
6: 2945257 Luc_Nilis
    [pem_pos=0, final_scores=0.6709314,
final_local_scores=0.7635176, pem=0.0, similarity=26.874601]
7: 1295837 Phillip_Cocu
    [pem_pos=0, final_scores=0.5629504,
final_local_scores=0.6019193, pem=-0.06525142,
similarity=22.407034]
```

Sample: b'56_SOCCER'

```
SOCCER - [1 1 DUTCH 1] FIRST DIVISION RESULT . [2 2 AMSTERDAM
2] 1996-08-22 Result of a [3 3 Dutch 3] first division match
on Thursday : [4 4 NEC Nijmegen 4] 1 [5 5 PSV Eindhoven 5] 4
1: 21148 Netherlands
2: 844 Amsterdam
3: 21148 Netherlands
4: 834154 N.E.C._(football_club)
5: 79820 PSV_Eindhoven
1: 21148 Netherlands
    [pem_pos=0, final_scores=0.68636936,
final_local_scores=0.66787416, pem=-0.7943452,
similarity=28.340183]
2: 844 Amsterdam
    [pem_pos=0, final_scores=0.63599175,
final_local_scores=0.63230956, pem=-0.0642803,
similarity=23.308659]
3: 21148 Netherlands
    [pem_pos=0, final_scores=0.68193626,
final_local_scores=0.66303164, pem=-0.7943452,
similarity=28.195673]
4: 834154 N.E.C._(football_club)
    [pem_pos=0, final_scores=0.54293036,
final_local_scores=0.5787057, pem=0.0, similarity=21.359446]
5: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.43539017,
final_local_scores=0.4415819, pem=0.0, similarity=17.267397]
```

Sample: b'57_SOCCER'

```
SOCCER - SHARPSHOOTER [1 1 KNUP 1] BACK IN [2 2 SWISS 2]
SQUAD . [3 3 GENEVA 3] 1996-08-22 [4 4 Galatasaray 4] striker
[5 5 Adrian Knup 5] , scorer of 26 goals in 45
internationals , has been recalled by [6 6 Switzerland 6] for
the [7 7 World Cup 7] qualifier against [8 8 Azerbaijan 8] in
[9 9 Baku 9] on August 31 . [10 10 Knup 10] was overlooked by
[11 11 Artur Jorge 11] for the [12 12 European 12]
```

championship finals earlier this year . But new coach [13 13 Rolf Fringer 13] is clearly a [14 14 Knup 14] fan and included him in his 19-man squad on Thursday . [15 15 Switzerland 15] failed to progress beyond the opening group phase in [16 16 Euro 96 16] . Squad : Goalkeepers - [17 17 Marco Pascolo 17] ([18 18 Cagliari 18]) , [19 19 Pascal Zuberbuehler 19] ([20 20 Grasshoppers 20]) . Defenders - [21 21 Stephane Henchoz 21] ([22 22 Hamburg 22]) , [23 23 Marc Hottiger 23] ([24 24 Everton 24]) , [25 25 Yvan Quentin 25] ([26 26 Sion 26]) , [27 27 Ramon Vega 27] ([28 28 Cagliari 28]) [29 29 Raphael Wicky 29] ([30 30 Sion 30]) . Midfielders - [31 31 Alexandre Camisetti 31] ([32 32 Grasshoppers 32]) , [33 33 Antonio Esposito 33] ([34 34 Grasshoppers 34]) , [35 35 Sebastien Fournier 35] ([36 36 Stuttgart 36]) , [37 37 Christophe Ohrel 37] ([38 38 Lausanne 38]) , [39 39 Patrick Sylvestre 39] ([40 40 Sion 40]) , [41 41 David Sesa 41] ([42 42 Servette 42]) , [43 43 Ciriaco Sforza 43] ([44 44 Inter Milan 44]) [45 45 Murat Yakin 45] ([46 46 Grasshoppers 46]) . Strikers - [47 47 Kubilay Turkyilmaz 47] ([48 48 Grasshoppers 48]) , [49 49 Adrian Knup 49] ([50 50 Galatasaray 50]) , [51 51 Christophe Bonvin 51] ([52 52 Sion 52]) , [53 53 Stephane Chauvet 53] ([54 54 Borussia Dortmund 54]) .

- 1: 5298672 Adrian_Knup
- 2: 26748 Switzerland
- 3: 12521 Geneva
- 4: 23608452 Galatasaray_S.K._(football)
- 5: 5298672 Adrian_Knup
- 6: 681862 Switzerland_national_football_team
- 7: 126460 1998_FIFA_World_Cup
- 8: 655906 Azerbaijan_national_football_team
- 9: 4566 Baku
- 10: 5298672 Adrian_Knup
- 11: 2695399 Artur_Jorge_(footballer)
- 12: 9239 Europe
- 13: 20120965 Rolf_Fringer
- 14: 5298672 Adrian_Knup
- 15: 681862 Switzerland_national_football_team
- 16: 555876 UEFA_Euro_1996
- 17: 5298205 Marco_Pascolo
- 18: 1116811 Cagliari_Calcio
- 19: 5160306 Pascal_Zuberbühler
- 20: 1573615 Grasshopper_Club_Zürich
- 21: 557033 Stéphane_Henchoz
- 22: 535879 Hamburger_SV
- 23: 5298238 Marc_Hottiger
- 24: 91155 Everton_F.C.
- 25: 5298407 Yvan_Quentin
- 26: 1061075 FC_Sion
- 27: 4596587 Ramon_Vega
- 28: 1116811 Cagliari_Calcio
- 29: 30874196 Raphaël_Wicky
- 30: 1061075 FC_Sion

```
31: 5728580 Alexandre_Comisetti
32: 1573615 Grasshopper_Club_Zürich
33: 19302946 Antonio_Esposito_(footballer_born_1972)
34: 1573615 Grasshopper_Club_Zürich
35: 5728499 Sébastien_Fournier
36: 737608 VfB_Stuttgart
37: 5298163 Christophe_Ohrel
38: 2092159 FC_Lausanne-Sport
39: 4993766 Patrick_Sylvestre
40: 1061075 FC_Sion
41: 5728607 David_Sesa
42: 893512 Servette_FC
43: 3921056 Ciriaco_Sforza
44: 15116 Inter_Milan
45: 383529 Murat_Yakin
46: 1573615 Grasshopper_Club_Zürich
47: 2645823 Kubilay_Türkyilmaz
48: 1573615 Grasshopper_Club_Zürich
49: 5298672 Adrian_Knup
50: 23608452 Galatasaray_S.K._(football)
51: 7377696 Christophe_Bonvin
52: 1061075 FC_Sion
53: 1287233 Stéphane_Chapuisat
54: 331715 Borussia_Dortmund
1: 5298672 Adrian_Knup
    [pem_pos=0, final_scores=0.54031044,
final_local_scores=0.57080257, pem=0.0, similarity=21.1236]
2: 26748 Switzerland
    [pem_pos=0, final_scores=0.3332162,
final_local_scores=0.36595917, pem=-0.1392413,
similarity=15.76788]
3: 12521 Geneva
    [pem_pos=0, final_scores=0.41562235,
final_local_scores=0.49150276, pem=-0.24772356,
similarity=20.104296]
4: 23608452 Galatasaray_S.K._(football)
    [pem_pos=1, final_scores=0.28739423,
final_local_scores=0.26800162, pem=-1.8990651,
similarity=22.414827]
5: 5298672 Adrian_Knup
    [pem_pos=0, final_scores=0.8858895,
final_local_scores=0.9482952, pem=0.0, similarity=32.388733]
6: 681862 Switzerland_national_football_team
    [pem_pos=1, final_scores=0.5948866,
final_local_scores=0.5732803, pem=-2.6520712,
similarity=35.619915]
7: 126460 1998_FIFA_World_Cup
    [pem_pos=18, final_scores=0.35014957,
final_local_scores=0.33433437, pem=-4.2172847,
similarity=37.001167]
8: 655906 Azerbaijan_national_football_team
    [pem_pos=1, final_scores=0.37529653,
```

```
final_local_scores=0.37444133, pem=-2.6132543,
similarity=29.47507]
9: 4566 Baku
    [pem_pos=0, final_scores=0.4250186,
final_local_scores=0.533228, pem=-0.06388413,
similarity=20.34971]
10: 5298672 Adrian_Knup
    [pem_pos=0, final_scores=0.7533507,
final_local_scores=0.8035166, pem=0.0, similarity=28.068253]
11: 2695399 Artur_Jorge_(footballer)
    [pem_pos=0, final_scores=0.5752444,
final_local_scores=0.63532597, pem=0.0, similarity=23.049107]
12: 9239 Europe
    [pem_pos=0, final_scores=0.544046,
final_local_scores=0.6795993, pem=-0.25557294,
similarity=25.760157]
13: 20120965 Rolf_Fringer
    [pem_pos=0, final_scores=0.712708,
final_local_scores=0.8092348, pem=0.0, similarity=28.238894]
14: 5298672 Adrian_Knup
    [pem_pos=0, final_scores=0.67229563,
final_local_scores=0.71497625, pem=0.0, similarity=25.426031]
15: 681862 Switzerland_national_football_team
    [pem_pos=1, final_scores=0.5414411,
final_local_scores=0.5138621, pem=-2.6520712,
similarity=33.846756]
16: 555876 UEFA_Euro_1996
    [pem_pos=0, final_scores=0.4683395,
final_local_scores=0.45487434, pem=0.0, similarity=17.66407]
17: 5298205 Marco_Pascolo
    [pem_pos=0, final_scores=0.71436024,
final_local_scores=0.7857242, pem=0.0, similarity=27.53729]
18: 1116811 Cagliari_Calcio
    [pem_pos=1, final_scores=0.3443768,
final_local_scores=0.3304752, pem=-0.8688384,
similarity=18.67663]
19: 5160306 Pascal_Zuberbühler
    [pem_pos=0, final_scores=0.9953204,
final_local_scores=1.0746837, pem=0.0, similarity=36.160416]
20: 1573615 Grasshopper_Club_Zürich
    [pem_pos=0, final_scores=0.6798976,
final_local_scores=0.65185976, pem=-0.21130908,
similarity=24.69164]
21: 557033 Stéphane_Henchoz
    [pem_pos=0, final_scores=0.8168145,
final_local_scores=0.91128784, pem=0.0, similarity=31.284359]
22: 535879 Hamburger_SV
    [pem_pos=1, final_scores=0.7210352,
final_local_scores=0.70963585, pem=-2.820485,
similarity=40.6049]
23: 5298238 Marc_Hottiger
    [pem_pos=0, final_scores=0.7403419,
```

```
final_local_scores=0.8260375, pem=0.0, similarity=28.74032]
24: 91155 Everton_F.C.
    [pem_pos=0, final_scores=0.42086083,
final_local_scores=0.41363478, pem=-0.17171584,
similarity=17.367218]
25: 5298407 Yvan_Quentin
    [pem_pos=0, final_scores=0.57474047,
final_local_scores=0.6813451, pem=0.0, similarity=24.42241]
26: 1061075 FC_Sion
    [pem_pos=1, final_scores=0.68208915,
final_local_scores=0.6702699, pem=-1.4381809,
similarity=31.912956]
27: 4596587 Ramon_Vega
    [pem_pos=0, final_scores=0.55902535,
final_local_scores=0.6029701, pem=0.0, similarity=22.083544]
28: 1116811 Cagliari_Calcio
    [pem_pos=1, final_scores=0.2936812,
final_local_scores=0.27509803, pem=-0.8688384,
similarity=17.024065]
29: 30874196 Raphaël_Wicky
    [pem_pos=0, final_scores=0.63528067,
final_local_scores=0.7275711, pem=0.0, similarity=25.801888]
30: 1061075 FC_Sion
    [pem_pos=1, final_scores=0.80500394,
final_local_scores=0.8045357, pem=-1.4381809,
similarity=35.919716]
31: 5728580 Alexandre_Comisetti
    [pem_pos=0, final_scores=0.5083139,
final_local_scores=0.6099161, pem=0.0, similarity=22.290825]
32: 1573615 Grasshopper_Club_Zürich
    [pem_pos=0, final_scores=0.576132,
final_local_scores=0.5385117, pem=-0.21130908,
similarity=21.309109]
33: 19302946 Antonio_Esposito_(footballer_born_1972)
    [pem_pos=2, final_scores=0.38152784,
final_local_scores=0.45725155, pem=-3.2580965,
similarity=35.453045]
34: 1573615 Grasshopper_Club_Zürich
    [pem_pos=0, final_scores=0.6361411,
final_local_scores=0.6040625, pem=-0.21130908,
similarity=23.265276]
35: 5728499 Sébastien_Fournier
    [pem_pos=0, final_scores=0.52151585,
final_local_scores=0.59007466, pem=0.0, similarity=21.698717]
36: 737608 VfB_Stuttgart
    [pem_pos=1, final_scores=0.66352314,
final_local_scores=0.6427315, pem=-2.1952448,
similarity=35.208183]
37: 5298163 Christophe_Ohrel
    [pem_pos=0, final_scores=0.76790553,
final_local_scores=0.8456439, pem=0.0, similarity=29.325413]
38: 2092159 FC_Lausanne-Sport
```

```
[pem_pos=3, final_scores=0.5417395,
final_local_scores=0.5149612, pem=-2.4908664,
similarity=33.0029]
39: 4993766 Patrick_Sylvestre
    [pem_pos=0, final_scores=0.6694462,
final_local_scores=0.7723736, pem=0.0, similarity=27.138882]
40: 1061075 FC_Sion
    [pem_pos=1, final_scores=0.8291369,
final_local_scores=0.8308973, pem=-1.4381809,
similarity=36.706398]
41: 5728607 David_Sesa
    [pem_pos=0, final_scores=0.64516896,
final_local_scores=0.74429953, pem=0.0, similarity=26.301096]
42: 893512 Servette_FC
    [pem_pos=0, final_scores=0.6112173,
final_local_scores=0.5948642, pem=0.0, similarity=21.841646]
43: 3921056 Ciriaco_Sforza
    [pem_pos=0, final_scores=0.6232794,
final_local_scores=0.6620714, pem=0.0, similarity=23.847244]
44: 15116 Inter_Milan
    [pem_pos=0, final_scores=0.36335284,
final_local_scores=0.32979095, pem=0.0, similarity=13.931333]
45: 383529 Murat_Yakin
    [pem_pos=0, final_scores=0.77894026,
final_local_scores=0.81477976, pem=0.0, similarity=28.404366]
46: 1573615 Grasshopper_Club_Zürich
    [pem_pos=0, final_scores=0.6416753,
final_local_scores=0.6101077, pem=-0.21130908,
similarity=23.445677]
47: 2645823 Kubilay_Türkyilmaz
    [pem_pos=0, final_scores=0.69045645,
final_local_scores=0.7703475, pem=0.0, similarity=27.078419]
48: 1573615 Grasshopper_Club_Zürich
    [pem_pos=0, final_scores=0.58263516,
final_local_scores=0.5456154, pem=-0.21130908,
similarity=21.521097]
49: 5298672 Adrian_Knup
    [pem_pos=0, final_scores=1.0187787,
final_local_scores=1.0934565, pem=0.0, similarity=36.720634]
50: 23608452 Galatasaray_S.K._(football)
    [pem_pos=1, final_scores=0.3252623,
final_local_scores=0.30936676, pem=-1.8990651,
similarity=23.649246]
51: 7377696 Christophe_Bonvin
    [pem_pos=0, final_scores=0.49742132,
final_local_scores=0.5931607, pem=0.0, similarity=21.790812]
52: 1061075 FC_Sion
    [pem_pos=1, final_scores=0.73715705,
final_local_scores=0.7304233, pem=-1.4381809,
similarity=33.708054]
53: 1287233 Stéphane_Chapuisat
    [pem_pos=0, final_scores=0.9804178,
```

```
final_local_scores=1.0381057, pem=0.0, similarity=35.068855]
54: 331715 Borussia_Dortmund
    [pem_pos=0, final_scores=0.44689578,
final_local_scores=0.4123773, pem=0.0, similarity=16.395874]
```

Sample: b'58_ATHLETICS'

ATHLETICS - IT 'S A RECORD - 40,000 BEERS ON THE HOUSE . [1 1
BRUSSELS 1] 1996-08-22 Spectators at Friday 's [2 2 Brussels 2
] grand prix meeting have an extra incentive to cheer on the
athletes to world record performances -- a free glass of
beer . A [3 3 Belgian 3] brewery has offered to pay for a
free round of drinks for all of the 40,000 crowd if a world
record goes at the meeting , organisers said on Thursday . It
could be one of the most expensive rounds of drinks ever .
The meeting is sold out already . Two world records are in
serious danger of being broken at the meeting -- the women 's
1,000 metres and the men 's 3,000 metres .

```
1: 3708 Brussels
2: 3708 Brussels
3: 3343 Belgium
1: 3708 Brussels
    [pem_pos=0, final_scores=0.6458508,
final_local_scores=0.58920497, pem=-0.025440974,
similarity=21.811117]
2: 3708 Brussels
    [pem_pos=0, final_scores=0.6204392,
final_local_scores=0.56144667, pem=-0.025440974,
similarity=20.982754]
3: 3343 Belgium
    [pem_pos=0, final_scores=0.6496745,
final_local_scores=0.62809414, pem=-0.12403572,
similarity=23.50782]
```

Sample: b'59_GOLF'

GOLF - [1 1 GERMAN OPEN 1] FIRST ROUND SCORES . [2 2 STUTTGART 2] , [3 3 Germany 3] 1996-08-22 Leading first round scores
in the [4 4 German Open 4] golf championship on Thursday ([5
5 Britain 5] unless stated) : 62 [6 6 Paul Broadhurst 6] 63 [
7 7 Raymond Russell 7] 64 [8 8 David J. Russell 8] , [9 9 Mic
hael Campbell 9] ([10 10 New Zealand 10]) , [11 11 Ian Woos
nam 11] , [12 12 Bernhard Langer 12] ([13 13 Germany 13]) ,
[14 14 Ronan Rafferty 14] , [15 15 Mats Lanner 15] ([16 16 Sw
eden 16]) , [17 17 Wayne Riley 17] ([18 18 Australia 18])
65 [19 19 Eamonn Darcy 19] ([20 20 Ireland 20]) , [21 21 Per
Nyman 21] ([22 22 Sweden 22]) , [23 23 Russell Claydon 23] ,
[24 24 Mark Roe 24] , [25 25 Retief Goosen 25] ([26 26 South
Africa 26]) , [27 27 Carl Suneson 27] 66 Stephen Field , [28
28 Paul Lawrie 28] , [29 29 Ian Pyman 29] , Max Anglert ([30
30 Sweden 30]) , [31 31 Miles Tunnicliff 31] , [32 32 Christ
ian Cevaer 32] ([33 33 France 33]) , [34 34 Des Smyth 34] ([35
35 Ireland 35]) , [36 36 David Carter 36] , [37 37 Lee We
stwood 37] , [38 38 Greg Chalmers 38] ([39 39 Australia

39]) , [40 40 Miguel_Angel_Martin 40] ([41 41 Spain 41]) ,
[42 42 Thomas_Bjorn 42] ([43 43 Denmark 43]) , Fernando
Roca ([44 44 Spain 44]) , [45 45 Derrick_Cooper 45] 67 Jeff
Hawksworth , [46 46 Padraig_Harrington 46] ([47 47 Ireland 47]
]) , Michael_Welch , Thomas_Gogele ([48 48 Germany 48]) , [
49 49 Paul_McGinley 49] ([50 50 Ireland 50]) , [51 51 Gary_0
rr 51] , [52 52 Jose-Maria_Canizares 52] ([53 53 Spain
53]) , [54 54 Michael_Jonzon 54] ([55 55 Sweden 55]) , [56
56 Paul_Eales 56] , David_Williams , [57 57 Andrew_Coltart 57]
, [58 58 Jonathan_Lomas 58] , [59 59 Jose_Rivero 59] ([60 60
Spain 60]) , [61 61 Robert_Karlsson 61] ([62 62 Sweden 62])
, Marcus_Wills , [63 63 Pedro_Linhart 63] ([64 64 Spain
64]) , [65 65 Jamie_Spence 65] , [66 66 Terry_Price 66] ([67
67 Australia 67]) , Juan_Carlos_Pinero ([68 68 Spain
68]) , [69 69 Mark_Mouland 69]
1: 4488686 German_Open_(golf)
2: 28565 Stuttgart
3: 11867 Germany
4: 4488686 German_Open_(golf)
5: 31717 United_Kingdom
6: 1688190 Paul_Broadhurst
7: 4375906 Raymond_Russell
8: 5340073 David_J_Russell
9: 2141373 Michael_Campbell
10: 4913064 New_Zealand
11: 1034991 Ian_Woosnam
12: 575815 Bernhard_Langer
13: 11867 Germany
14: 1519060 Ronan_Rafferty
15: 4919505 Mats_Lanner
16: 5058739 Sweden
17: 5243170 Wayne_Riley
18: 4689264 Australia
19: 2770292 Eamonn_Darcy
20: 14560 Republic_of_Ireland
21: 28343641 Per_Nyman
22: 5058739 Sweden
23: 4291546 Russell_Claydon
24: 4353394 Mark_Roe
25: 630525 Retief_Goosen
26: 17416221 South_Africa
27: 11818919 Carl_Suneson
28: 1564436 Paul_Lawrie
29: 5088382 Iain_Pyman
30: 5058739 Sweden
31: 4347331 Miles_Tunnicliff
32: 4380538 Christian_Cévaër
33: 5843419 France
34: 5193762 Des_Smyth
35: 14560 Republic_of_Ireland
36: 4344819 David_Carter_(golfer)
37: 667682 Lee_Westwood

```
38: 5243345 Greg_Chalmers
39: 4689264 Australia
40: 5022198 Miguel_Ángel_Martín_(golfer)
41: 26667 Spain
42: 1891067 Thomas_Bjørn
43: 76972 Denmark
44: 26667 Spain
45: 5555797 Derrick_Cooper
46: 1157761 Pádraig_Harrington
47: 14560 Republic_of_Ireland
48: 11867 Germany
49: 1342124 Paul_McGinley
50: 14560 Republic_of_Ireland
51: 2167716 Gary_Orr
52: 4455122 José_Maria_Cañizares
53: 26667 Spain
54: 5242970 Michael_Jonzon
55: 5058739 Sweden
56: 4435993 Paul_Eales
57: 2150917 Andrew_Coltart
58: 5272617 Jonathan_Lomas
59: 4996266 José_Rivero
60: 26667 Spain
61: 1988802 Robert_Karlsson
62: 5058739 Sweden
63: 5285076 Pedro_Linhart
64: 26667 Spain
65: 3535223 Jamie_Spence
66: 22856557 Terry_Price_(golfer)
67: 4689264 Australia
68: 26667 Spain
69: 4384175 Mark_Mouland
1: 4488686 German_Open_(golf)
    [pem_pos=0, final_scores=0.5984021,
final_local_scores=0.6708363, pem=-0.62093085,
similarity=27.485525]
2: 28565 Stuttgart
    [pem_pos=0, final_scores=0.52052665,
final_local_scores=0.5868042, pem=-0.15141529,
similarity=22.424541]
3: 11867 Germany
    [pem_pos=0, final_scores=0.74699444,
final_local_scores=0.8057524, pem=-0.15529114,
similarity=28.97947]
4: 4488686 German_Open_(golf)
    [pem_pos=0, final_scores=0.44166374,
final_local_scores=0.49962348, pem=-0.62093085,
similarity=22.376192]
5: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5392906,
final_local_scores=0.58865017, pem=-0.61114657,
similarity=24.97972]
```

```
6: 1688190 Paul_Broadhurst
    [pem_pos=0, final_scores=0.9181487,
final_local_scores=1.0060863, pem=0.0, similarity=34.113335]
7: 4375906 Raymond_Russell
    [pem_pos=0, final_scores=0.6262997,
final_local_scores=0.6795699, pem=0.0, similarity=24.369434]
8: 5340073 David_J_Russell
    [pem_pos=0, final_scores=0.5000503,
final_local_scores=0.55500984, pem=0.0, similarity=20.652313]
9: 2141373 Michael_Campbell
    [pem_pos=0, final_scores=0.74496156,
final_local_scores=0.79716057, pem=0.0, similarity=27.878574]
10: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.43560815,
final_local_scores=0.48224598, pem=-0.23795462,
similarity=19.77493]
11: 1034991 Ian_Woosnam
    [pem_pos=0, final_scores=0.79546374,
final_local_scores=0.84540737, pem=0.0, similarity=29.318356]
12: 575815 Bernhard_Langer
    [pem_pos=0, final_scores=0.847725,
final_local_scores=0.90161806, pem=0.0, similarity=30.995794]
13: 11867 Germany
    [pem_pos=0, final_scores=0.7962652,
final_local_scores=0.8595732, pem=-0.15529114,
similarity=30.585588]
14: 1519060 Ronan_Rafferty
    [pem_pos=0, final_scores=0.5503035,
final_local_scores=0.59424114, pem=0.0, similarity=21.823053]
15: 4919505 Mats_Lanner
    [pem_pos=0, final_scores=0.7242525,
final_local_scores=0.7841704, pem=0.0, similarity=27.490921]
16: 5058739 Sweden
    [pem_pos=0, final_scores=0.70896626,
final_local_scores=0.7453871, pem=-0.2135508,
similarity=27.494871]
17: 5243170 Wayne_Riley
    [pem_pos=0, final_scores=0.72959226,
final_local_scores=0.78767425, pem=0.0, similarity=27.595484]
18: 4689264 Australia
    [pem_pos=0, final_scores=0.7365251,
final_local_scores=0.8028831, pem=-0.17168477,
similarity=28.982994]
19: 2770292 Eamonn_Darcy
    [pem_pos=0, final_scores=0.55663633,
final_local_scores=0.6130956, pem=0.0, similarity=22.385708]
20: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.5644818,
final_local_scores=0.61183506, pem=-1.5040284,
similarity=30.527231]
21: 28343641 Per_Nyman
    [pem_pos=0, final_scores=0.5994914,
```

```
final_local_scores=0.68801814, pem=0.0, similarity=24.621546]
22: 5058739 Sweden
    [pem_pos=0, final_scores=0.7325461,
final_local_scores=0.77114445, pem=-0.2135508,
similarity=28.263523]
23: 4291546 Russell_Claydon
    [pem_pos=0, final_scores=0.8161287,
final_local_scores=0.89713556, pem=0.0, similarity=30.862028]
24: 4353394 Mark_Roe
    [pem_pos=0, final_scores=0.7971646,
final_local_scores=0.88550454, pem=0.0, similarity=30.514935]
25: 630525 Retief_Goosen
    [pem_pos=0, final_scores=0.77455753,
final_local_scores=0.85406786, pem=0.0, similarity=29.576801]
26: 17416221 South_Africa
    [pem_pos=0, final_scores=0.5911394,
final_local_scores=0.6616515, pem=-0.26251382,
similarity=25.262304]
27: 11818919 Carl_Suneson
    [pem_pos=0, final_scores=0.82593924,
final_local_scores=0.9025677, pem=0.0, similarity=31.024132]
28: 1564436 Paul_Lawrie
    [pem_pos=0, final_scores=0.764239,
final_local_scores=0.83204794, pem=0.0, similarity=28.919683]
29: 5088382 Iain_Pyman
    [pem_pos=0, final_scores=0.7487145,
final_local_scores=0.8300276, pem=0.0, similarity=28.859392]
30: 5058739 Sweden
    [pem_pos=0, final_scores=0.7482794,
final_local_scores=0.7883307, pem=-0.2135508,
similarity=28.776394]
31: 4347331 Miles_Tunnicliff
    [pem_pos=0, final_scores=0.7911796,
final_local_scores=0.88364404, pem=0.0, similarity=30.459415]
32: 4380538 Christian_Cévaër
    [pem_pos=0, final_scores=0.74158365,
final_local_scores=0.81048536, pem=0.0, similarity=28.276213]
33: 5843419 France
    [pem_pos=0, final_scores=0.68443614,
final_local_scores=0.73935837, pem=-0.16240834,
similarity=27.036842]
34: 5193762 Des_Smyth
    [pem_pos=0, final_scores=0.7966898,
final_local_scores=0.8698296, pem=0.0, similarity=30.047163]
35: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.59985644,
final_local_scores=0.6504764, pem=-1.5040284,
similarity=31.680367]
36: 4344819 David_Carter_(golfer)
    [pem_pos=4, final_scores=0.64426786,
final_local_scores=0.69717723, pem=-1.8641758,
similarity=35.03255]
```

```
37: 667682 Lee_Westwood
    [pem_pos=0, final_scores=0.61979234,
final_local_scores=0.6563206, pem=0.0, similarity=23.675629]
38: 5243345 Greg_Chalmers
    [pem_pos=0, final_scores=0.74757135,
final_local_scores=0.81299794, pem=0.0, similarity=28.351192]
39: 4689264 Australia
    [pem_pos=0, final_scores=0.64908224,
final_local_scores=0.707365, pem=-0.17168477,
similarity=26.132544]
40: 5022198 Miguel_Ángel_Martín_(golfer)
    [pem_pos=0, final_scores=0.6983634,
final_local_scores=0.76614785, pem=0.0, similarity=26.953093]
41: 26667 Spain
    [pem_pos=0, final_scores=0.6059548,
final_local_scores=0.639572, pem=-0.18857051,
similarity=24.201294]
42: 1891067 Thomas_Bjørn
    [pem_pos=0, final_scores=0.7963721,
final_local_scores=0.8553547, pem=0.0, similarity=29.615204]
43: 76972 Denmark
    [pem_pos=0, final_scores=0.6402847,
final_local_scores=0.6813553, pem=-0.26762357,
similarity=25.878092]
44: 26667 Spain
    [pem_pos=0, final_scores=0.6109682,
final_local_scores=0.64504826, pem=-0.18857051,
similarity=24.364716]
45: 5555797 Derrick_Cooper
    [pem_pos=0, final_scores=0.7030869,
final_local_scores=0.7828075, pem=0.0, similarity=27.45025]
46: 1157761 Pádraig_Harrington
    [pem_pos=0, final_scores=0.65844256,
final_local_scores=0.70171756, pem=0.0, similarity=25.030365]
47: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.5818297,
final_local_scores=0.63078505, pem=-1.5040284,
similarity=31.09274]
48: 11867 Germany
    [pem_pos=0, final_scores=0.79675657,
final_local_scores=0.86010987, pem=-0.15529114,
similarity=30.601606]
49: 1342124 Paul_McGinley
    [pem_pos=0, final_scores=0.712215,
final_local_scores=0.7537579, pem=0.0, similarity=26.583351]
50: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.55851,
final_local_scores=0.6053118, pem=-1.5040284,
similarity=30.332567]
51: 2167716 Gary_Orr
    [pem_pos=0, final_scores=0.58631843,
final_local_scores=0.6601269, pem=0.0, similarity=23.789217]
```

52: 4455122 José_Maria_Cañizares
[pem_pos=0, final_scores=0.7736899,
final_local_scores=0.84818363, pem=0.0, similarity=29.401205]
53: 26667 Spain
[pem_pos=0, final_scores=0.62728804,
final_local_scores=0.66287524, pem=-0.18857051,
similarity=24.896708]
54: 5242970 Michael_Jonzon
[pem_pos=0, final_scores=0.66900605,
final_local_scores=0.7490772, pem=0.0, similarity=26.44367]
55: 5058739 Sweden
[pem_pos=0, final_scores=0.7101683,
final_local_scores=0.7467001, pem=-0.2135508,
similarity=27.534056]
56: 4435993 Paul_Eales
[pem_pos=0, final_scores=0.69861704,
final_local_scores=0.77122295, pem=0.0, similarity=27.104544]
57: 2150917 Andrew_Coltart
[pem_pos=0, final_scores=0.71306664,
final_local_scores=0.7607561, pem=0.0, similarity=26.792192]
58: 5272617 Jonathan_Lomas
[pem_pos=0, final_scores=0.6505828,
final_local_scores=0.7243707, pem=0.0, similarity=25.70638]
59: 4996266 José_Rivero
[pem_pos=0, final_scores=0.46924865,
final_local_scores=0.5101159, pem=0.0, similarity=19.31259]
60: 26667 Spain
[pem_pos=0, final_scores=0.637297,
final_local_scores=0.6738086, pem=-0.18857051,
similarity=25.22298]
61: 1988802 Robert_Karlsson
[pem_pos=0, final_scores=0.78203684,
final_local_scores=0.86536187, pem=0.0, similarity=29.913837]
62: 5058739 Sweden
[pem_pos=0, final_scores=0.7389878,
final_local_scores=0.7781811, pem=-0.2135508,
similarity=28.47351]
63: 5285076 Pedro_Linhart
[pem_pos=0, final_scores=0.7844345,
final_local_scores=0.8742854, pem=0.0, similarity=30.180132]
64: 26667 Spain
[pem_pos=0, final_scores=0.634646,
final_local_scores=0.67091274, pem=-0.18857051,
similarity=25.136562]
65: 3535223 Jamie_Spence
[pem_pos=0, final_scores=0.48222196,
final_local_scores=0.59284025, pem=0.0, similarity=21.781248]
66: 22856557 Terry_Price_(golfer)
[pem_pos=1, final_scores=0.56414163,
final_local_scores=0.64302665, pem=-0.47743857,
similarity=25.875296]
67: 4689264 Australia

```
[pem_pos=0, final_scores=0.6382714,
final_local_scores=0.69555575, pem=-0.17168477,
similarity=25.780132]
68: 26667 Spain
    [pem_pos=0, final_scores=0.6191887,
final_local_scores=0.654028, pem=-0.18857051,
similarity=24.632689]
69: 4384175 Mark Moulard
    [pem_pos=0, final_scores=0.59762484,
final_local_scores=0.65481794, pem=0.0, similarity=23.630787]
```

Sample: b'60_SOCCER'

SOCCER - [1 1 UEFA 1] REWARDS THREE COUNTRIES FOR FAIR PLAY . [2 2 GENEVA 2] 1996-08-22 [3 3 Norway 3] , [4 4 England 4] and [5 5 Sweden 5] were rewarded for their fair play on Thursday with an additional place in the 1997-98 [6 6 UEFA Cup 6] competition . [7 7 Norway 7] headed the UEFA Fair Play rankings for 1995-96 with 8.62 points , ahead of [8 8 England 8] with 8.61 and [9 9 Sweden 9] 8.57 . The rankings are based on a formula that takes into account many factors including red and yellow cards , and coaching and spectators ' behaviour at matches played at an international level by clubs and national teams . Only the top three countries are allocated additional places . The UEFA Fair Play rankings are : 1 . [10 10 Norway 10] 8.62 points 2. [11 11 England 11] 8.61 3. [12 12 Sweden 12] 8.57 4. [13 13 Faroe Islands 13] 8.56 5. [14 14 Wales 14] 8.54 6. [15 15 Estonia 15] 8.52 7. [1 6 16 Ireland 16] 8.45 8. [17 17 Belarus 17] 8.39 9. [18 18 Iceland 18] 8.35 10. [19 19 Netherlands 19] 8.30 10. [20 20 Denmark 20] 8.30 10. [21 21 Germany 21] 8.30 13. [22 22 Scotland 22] 8.29 13. [23 23 Latvia 23] 8.29 15. [24 24 Moldova 24] 8.24 16. [25 25 Yugoslavia 25] 8.22 16. [26 26 Belgium 26] 8.22 18. [27 27 Luxembourg 27] 8.20 19. [28 28 France 28] 8.18 20. [29 29 Israel 29] 8.17 21. [30 30 Switzerland 30] 8.15 21. [31 31 Slovakia 31] 8.15 23. [32 32 Poland 32] 8.12 23. [33 33 Portugal 33] 8.12 25. [34 34 Georgia 34] 8.10 26. [35 35 Ukraine 35] 8.09 26. [36 36 Spain 36] 8.09 26. [37 37 Finland 37] 8.09 29. [38 38 Macedonia 38] 8.07 30. [39 39 Lithuania 39] 8.06 31. [40 40 Austria 40] 8.05 32. [41 41 Russia 41] 8.03 33. [42 42 Romania 42] 8.02 33. Turkey 8.02 35. [43 43 Hungary 43] 7.98 36. [44 44 Czech Republic 44] 7.95 37. [45 45 Greece 45] 7.89 37. [46 46 Northern Ireland 46] 7.89 39. [47 47 Italy 47] 7.85 40. [48 48 Cyprus 48] 7.83 41. [49 49 Armenia 49] 7.80 42. [50 50 Slovenia 50] 7.77 43. [51 51 Croatia 51] 7.75 44. [52 52 Bulgaria 52] 7.73 45. [53 53 Malta 53] 7.40

1: 32332 UEFA
2: 12521 Geneva
3: 21241 Norway
4: 9316 England
5: 5058739 Sweden
6: 232175 UEFA_Europa_League
7: 21241 Norway

8: 9316 England
9: 5058739 Sweden
10: 21241 Norway
11: 9316 England
12: 5058739 Sweden
13: 17238607 Faroe_Islands
14: 69894 Wales
15: 28222445 Estonia
16: 14560 Republic_of_Ireland
17: 3457 Belarus
18: 14531 Iceland
19: 21148 Netherlands
20: 76972 Denmark
21: 11867 Germany
22: 26994 Scotland
23: 17514 Latvia
24: 19260 Moldova
25: 34244 Yugoslavia
26: 3343 Belgium
27: 17515 Luxembourg
28: 5843419 France
29: 9282173 Israel
30: 26748 Switzerland
31: 26830 Slovakia
32: 22936 Poland
33: 23033 Portugal
34: 48768 Georgia_(country)
35: 31750 Ukraine
36: 26667 Spain
37: 10577 Finland
38: 23564616 Republic_of_Macedonia
39: 17675 Lithuania
40: 26964606 Austria
41: 25391 Russia
42: 25445 Romania
43: 13275 Hungary
44: 5321 Czech_Republic
45: 12108 Greece
46: 21265 Northern_Ireland
47: 14532 Italy
48: 5593 Cyprus
49: 10918072 Armenia
50: 27338 Slovenia
51: 5573 Croatia
52: 3415 Bulgaria
53: 19137 Malta
1: 32332 UEFA
 [pem_pos=0, final_scores=0.28276944,
final_local_scores=0.30977485, pem=0.0, similarity=13.334012]
2: 12521 Geneva
 [pem_pos=0, final_scores=0.47650605,
final_local_scores=0.4925692, pem=-0.24772356,

```
similarity=20.13612]
3: 21241 Norway
    [pem_pos=0, final_scores=0.6861194,
final_local_scores=0.67642486, pem=-0.22668563,
similarity=25.508331]
4: 9316 England
    [pem_pos=0, final_scores=0.40342182,
final_local_scores=0.3892541, pem=-0.19053856,
similarity=16.74201]
5: 5058739 Sweden
    [pem_pos=0, final_scores=0.5606008,
final_local_scores=0.5192535, pem=-0.2135508,
similarity=20.746595]
6: 232175 UEFA_Europa_League
    [pem_pos=0, final_scores=0.22184534,
final_local_scores=0.293351, pem=-0.260892,
similarity=14.262663]
7: 21241 Norway
    [pem_pos=0, final_scores=0.55877966,
final_local_scores=0.5373255, pem=-0.22668563,
similarity=21.35733]
8: 9316 England
    [pem_pos=0, final_scores=0.424223,
final_local_scores=0.41197622, pem=-0.19053856,
similarity=17.420084]
9: 5058739 Sweden
    [pem_pos=0, final_scores=0.558611,
final_local_scores=0.51707983, pem=-0.2135508,
similarity=20.681728]
10: 21241 Norway
    [pem_pos=0, final_scores=0.52863026,
final_local_scores=0.50439185, pem=-0.22668563,
similarity=20.374523]
11: 9316 England
    [pem_pos=0, final_scores=0.4693907,
final_local_scores=0.46131498, pem=-0.19053856,
similarity=18.89245]
12: 5058739 Sweden
    [pem_pos=0, final_scores=0.78297406,
final_local_scores=0.7621624, pem=-0.2135508,
similarity=27.995481]
13: 17238607 Faroe_Islands
    [pem_pos=0, final_scores=0.3569966,
final_local_scores=0.3384474, pem=-0.59686494,
similarity=17.435503]
14: 69894 Wales
    [pem_pos=0, final_scores=0.45896846,
final_local_scores=0.46293294, pem=-0.37222588,
similarity=19.928778]
15: 28222445 Estonia
    [pem_pos=0, final_scores=0.74043804,
final_local_scores=0.73480725, pem=-0.4106629,
```

```
similarity=28.251076]
16: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.38204348,
final_local_scores=0.3716815, pem=-1.5040284,
similarity=23.36057]
17: 3457 Belarus
    [pem_pos=0, final_scores=0.69975835,
final_local_scores=0.68286216, pem=-0.4220667,
similarity=26.762945]
18: 14531 Iceland
    [pem_pos=0, final_scores=0.60737056,
final_local_scores=0.5908935, pem=-0.36386928,
similarity=23.701931]
19: 21148 Netherlands
    [pem_pos=0, final_scores=0.62996143,
final_local_scores=0.6128009, pem=-0.24940795,
similarity=23.733234]
20: 76972 Denmark
    [pem_pos=0, final_scores=0.6569418,
final_local_scores=0.628113, pem=-0.26762357,
similarity=24.289236]
21: 11867 Germany
    [pem_pos=0, final_scores=0.95830566,
final_local_scores=0.9558825, pem=-0.15529114,
similarity=33.459652]
22: 26994 Scotland
    [pem_pos=0, final_scores=0.43952376,
final_local_scores=0.43599725, pem=-0.22905664,
similarity=18.346386]
23: 17514 Latvia
    [pem_pos=0, final_scores=0.79211336,
final_local_scores=0.7751316, pem=-0.35144013,
similarity=29.132372]
24: 19260 Moldova
    [pem_pos=0, final_scores=0.54846936,
final_local_scores=0.52446944, pem=-0.47343826,
similarity=22.31556]
25: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.4140817,
final_local_scores=0.3764757, pem=-0.6536551,
similarity=18.879177]
26: 3343 Belgium
    [pem_pos=0, final_scores=0.7688218,
final_local_scores=0.7551073, pem=-0.24180195,
similarity=27.938578]
27: 17515 Luxembourg
    [pem_pos=0, final_scores=0.5584567,
final_local_scores=0.5262464, pem=-0.49845144,
similarity=22.504612]
28: 5843419 France
    [pem_pos=0, final_scores=0.75476736,
final_local_scores=0.7466736, pem=-0.16240834,
```

```
similarity=27.255142]
29: 9282173 Israel
    [pem_pos=0, final_scores=0.61646086,
final_local_scores=0.6300579, pem=-0.21279564,
similarity=24.049112]
30: 26748 Switzerland
    [pem_pos=0, final_scores=0.71830136,
final_local_scores=0.69835067, pem=-0.18598008,
similarity=25.941278]
31: 26830 Slovakia
    [pem_pos=0, final_scores=0.6837731,
final_local_scores=0.65646565, pem=-0.36602598,
similarity=25.670464]
32: 22936 Poland
    [pem_pos=0, final_scores=0.7602224,
final_local_scores=0.7352771, pem=-0.15333635,
similarity=26.865715]
33: 23033 Portugal
    [pem_pos=0, final_scores=0.59969914,
final_local_scores=0.58523446, pem=-0.30942732,
similarity=23.23699]
34: 48768 Georgia_(country)
    [pem_pos=1, final_scores=0.38313544,
final_local_scores=0.35254478, pem=-1.3466994,
similarity=21.933912]
35: 31750 Ukraine
    [pem_pos=0, final_scores=0.6661463,
final_local_scores=0.64266074, pem=-0.35407752,
similarity=25.19352]
36: 26667 Spain
    [pem_pos=0, final_scores=0.6446262,
final_local_scores=0.6326247, pem=-0.18857051,
similarity=23.99397]
37: 10577 Finland
    [pem_pos=0, final_scores=0.81279093,
final_local_scores=0.8007084, pem=-0.2445894,
similarity=29.314564]
38: 23564616 Republic_of_Macedonia
    [pem_pos=0, final_scores=0.5150556,
final_local_scores=0.4822743, pem=-1.1139666,
similarity=24.539663]
39: 17675 Lithuania
    [pem_pos=0, final_scores=0.67854506,
final_local_scores=0.6495136, pem=-0.39766282,
similarity=25.635046]
40: 26964606 Austria
    [pem_pos=0, final_scores=0.61788744,
final_local_scores=0.5788188, pem=-0.25262418,
similarity=22.73663]
41: 25391 Russia
    [pem_pos=0, final_scores=0.79196197,
final_local_scores=0.80136794, pem=-0.20299491,
```

```
similarity=29.108047]
42: 25445 Romania
    [pem_pos=0, final_scores=0.74876136,
final_local_scores=0.7335587, pem=-0.21241479,
similarity=27.135712]
43: 13275 Hungary
    [pem_pos=0, final_scores=0.7111241,
final_local_scores=0.68507624, pem=-0.32577527,
similarity=26.305372]
44: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.57148725,
final_local_scores=0.53325427, pem=-0.27794555,
similarity=21.514595]
45: 12108 Greece
    [pem_pos=0, final_scores=0.95963544,
final_local_scores=0.9712172, pem=-0.27607638,
similarity=34.57412]
46: 21265 Northern_Ireland
    [pem_pos=0, final_scores=0.43974334,
final_local_scores=0.4390449, pem=-0.25044483,
similarity=18.553646]
47: 14532 Italy
    [pem_pos=0, final_scores=0.77049726,
final_local_scores=0.7748461, pem=-0.20612453,
similarity=28.333603]
48: 5593 Cyprus
    [pem_pos=0, final_scores=0.66984636,
final_local_scores=0.66440016, pem=-0.3752463,
similarity=25.957386]
49: 10918072 Armenia
    [pem_pos=0, final_scores=0.5944106,
final_local_scores=0.59014964, pem=-0.4307794,
similarity=24.0436]
50: 27338 Slovenia
    [pem_pos=0, final_scores=0.41891032,
final_local_scores=0.39116287, pem=-0.27845925,
similarity=17.277098]
51: 5573 Croatia
    [pem_pos=0, final_scores=0.66966754,
final_local_scores=0.6428196, pem=-0.35021794,
similarity=25.17727]
52: 3415 Bulgaria
    [pem_pos=0, final_scores=0.802098,
final_local_scores=0.7933198, pem=-0.31727514,
similarity=29.489351]
53: 19137 Malta
    [pem_pos=0, final_scores=0.55492896,
final_local_scores=0.55259675, pem=-0.39222723,
similarity=22.713295]

Sample: b'61_CRICKET'
CRICKET - POLICE COMMANDOS ON HAND FOR [1 1 AUSTRALIANS 1] '
```

FIRST MATCH . [2 2 COLOMBO 2] 1996-08-22 Armed police commandos patrolled the ground when [3 3 Australia 3] opened their short tour of [4 4 Sri Lanka 4] with a five-run win over the country 's youth team on Thursday . [5 5 Australia 5] , in [6 6 Sri Lanka 6] for a limited overs tournament which also includes [7 7 India 7] and [8 8 Zimbabwe 8] , have been promised the presence of commandos , sniffer dogs and plainclothes policemen to ensure the tournament is trouble-free . They are making their first visit to the island since boycotting a [1 World Cup 1] fixture in February because of fears over ethnic violence . [9 9 Australia 9] , batting first in Thursday 's the warm-up match , scored 251 for seven from their 50 overs . [10 10 Ricky Ponting 10] led the way with 100 off 119 balls with two sixes and nine fours before retiring . The youth side replied with 246 for seven . [11 11 Australian 11] coach [12 12 Geoff Marsh 12] said he was impressed with the competitiveness of the opposition . " We were made to sweat to win , " he said .

1: 374833 Australia_national_cricket_team
2: 56636 Colombo
3: 374833 Australia_national_cricket_team
4: 26750 Sri_Lanka
5: 374833 Australia_national_cricket_team
6: 26750 Sri_Lanka
7: 407754 India_national_cricket_team
8: 759317 Zimbabwe_national_cricket_team
9: 374833 Australia_national_cricket_team
10: 462001 Ricky_Ponting
11: 374833 Australia_national_cricket_team
12: 1516661 Geoff_Marsh
1: 374833 Australia_national_cricket_team
[pem_pos=2, final_scores=0.54008096,
final_local_scores=0.47875756, pem=-1.4598138,
similarity=26.315487]
2: 56636 Colombo
[pem_pos=0, final_scores=0.5683071,
final_local_scores=0.6044904, pem=-0.19393882,
similarity=23.183582]
3: 374833 Australia_national_cricket_team
[pem_pos=1, final_scores=0.6940684,
final_local_scores=0.6469654, pem=-3.4989398,
similarity=42.42423]
4: 26750 Sri_Lanka
[pem_pos=0, final_scores=0.46760207,
final_local_scores=0.50699365, pem=-0.3027737,
similarity=20.865946]
5: 374833 Australia_national_cricket_team
[pem_pos=1, final_scores=0.78606564,
final_local_scores=0.74745846, pem=-3.4989398,
similarity=45.42314]
6: 26750 Sri_Lanka
[pem_pos=0, final_scores=0.41507596,

```
final_local_scores=0.4496168, pem=-0.3027737,
similarity=19.153706]
7: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.6958948,
final_local_scores=0.65847003, pem=-3.5152702,
similarity=42.856358]
8: 759317 Zimbabwe_national_cricket_team
    [pem_pos=1, final_scores=0.91354054,
final_local_scores=0.91228205, pem=-2.2776897,
similarity=43.70046]
9: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.7751009,
final_local_scores=0.73548114, pem=-3.4989398,
similarity=45.065716]
10: 462001 Ricky_Ponting
    [pem_pos=0, final_scores=0.3923738,
final_local_scores=0.37990403, pem=0.0, similarity=15.426807]
11: 374833 Australia_national_cricket_team
    [pem_pos=2, final_scores=0.6693509,
final_local_scores=0.61996543, pem=-3.6253178,
similarity=42.30576]
12: 1516661 Geoff_Marsh
    [pem_pos=0, final_scores=0.78964454,
final_local_scores=0.834063, pem=0.0, similarity=28.979816]
1: 7239 Cricket_World_Cup
    [pem_pos=4, final_scores=0.21361987,
final_local_scores=0.17675447, pem=-3.1731894,
similarity=26.620716]
```

Sample: b'62_ONE'

ONE [1 1 ROMANIAN 1] DIES IN BUS CRASH IN [2 2 BULGARIA 2] .
[3 3 SOFIA 3] 1996-08-22 One [4 4 Romanian 4] passenger was killed , and 14 others were injured on Thursday when a Romanian-registered bus collided with a [5 5 Bulgarian 5] one in northern [6 6 Bulgaria 6] , police said . The two buses collided head on at 5 o'clock this morning on the road between the towns of [7 7 Rousse 7] and [8 8 Veliko Tarnovo 8] , police said . A [9 9 Romanian 9] woman Maria Marco , 35 , was killed . The accident was being investigated , police added . -- Sofia Newsroom , 359-2-84561

```
1: 25445 Romania
2: 3415 Bulgaria
3: 57644 Sofia
4: 25445 Romania
5: 3415 Bulgaria
6: 3415 Bulgaria
7: 274597 Ruse,_Bulgaria
8: 193500 Veliko_Tarnovo
9: 25445 Romania
1: 25445 Romania
    [pem_pos=1, final_scores=0.676873,
final_local_scores=0.6482967, pem=-1.0382323,
```

```
similarity=29.082249]
2: 3415 Bulgaria
    [pem_pos=0, final_scores=0.54417896,
final_local_scores=0.489555, pem=-0.31727514,
similarity=20.424402]
3: 57644 Sofia
    [pem_pos=0, final_scores=0.57051617,
final_local_scores=0.54224783, pem=-0.10680699,
similarity=20.852303]
4: 25445 Romania
    [pem_pos=1, final_scores=0.72435814,
final_local_scores=0.70016694, pem=-1.0382323,
similarity=30.63016]
5: 3415 Bulgaria
    [pem_pos=1, final_scores=0.53821,
final_local_scores=0.48303485, pem=-1.1560022,
similarity=24.790956]
6: 3415 Bulgaria
    [pem_pos=0, final_scores=0.71225446,
final_local_scores=0.673152, pem=-0.31727514,
similarity=25.903303]
7: 274597 Ruse,_Bulgaria
    [pem_pos=0, final_scores=0.42789465,
final_local_scores=0.42198956, pem=0.0, similarity=16.682724]
8: 193500 Veliko_Tarnovo
    [pem_pos=0, final_scores=0.36443442,
final_local_scores=0.35315382, pem=0.0, similarity=14.628528]
9: 25445 Romania
    [pem_pos=1, final_scores=0.70635843,
final_local_scores=0.68050504, pem=-1.0382323,
similarity=30.043411]
```

Sample: b'63_OFFICIAL'

OFFICIAL JOURNAL CONTENTS - OJ L 211 OF AUGUST 21 , 1996 . *

(Note - contents are displayed in reverse order to that in the printed Journal) * Corrigendum to Commission Regulation (EC) No 1464/96 of 25 July 1996 relating to a standing invitation to tender to determine levies and / or refunds on exports of white sugar (OJ No L 187 of 26.7.1996)

Corrigendum to Commission Regulation (EC) No 658/96 of 9 April 1996 on certain conditions for granting compensatory payments under the support system for producers of certain arable crops (OJ No L 91 of 12.4.1996)

Commission Regulation (EC) No 1663/96 of 20 August 1996 establishing the standard import values for determining the entry price of certain fruit and vegetables END OF DOCUMENT .

Sample: b'64_In'

In Home Health to appeal payment denial . [1 1 MINNETONKA 1] , [2 2 Minn 2] . 1996-08-22 In Home Health Inc said on Thursday it will appeal to the [3 3 U.S. Federal District Cou

rt 3] in [4 4 Minneapolis 4] a decision by the Health Care Financing Administration (HCFA) that denied reimbursement of certain costs under Medicaid . The HCFA Administrator reversed a previously favorable decision regarding the reimbursement of costs related to the company 's community liaison personnel , it added . The company said it continues to believe the majority of the community liaison costs are coverable under the terms of the [5 5 Medicare 5] program . " We are disappointed with the administrator 's decision but we continue to be optimistic regarding an ultimate favorable resolution , " Mark Gildea , chief executive officer , said in a statement . In Home Health said it previously recorded a reserve equal to 16 percent of all revenue related to the community liaison costs . Separately , In Home Health said the [6 6 U.S. District Court 6] in [7 7 Minneapolis 7] ruled in its favor regarding the reimbursement of certain interest expenses . This decision will result in the reimbursement by [8 8 Medicare 8] of \$ 81,000 in disputed costs . " This is our first decision in federal district court regarding a dispute with [9 9 Medicare 9] , " Gildea said . " We are extremely pleased with this decision and we recognize it as a significant step toward resolution of our outstanding [10 10 Medicare 10] disputes . " -- Chicago Newsdesk 312-408-8787

1: 120211 Minnetonka,_Minnesota
2: 19590 Minnesota
3: 70235 United_States_district_court
4: 6097240 Minneapolis
5: 58353 Medicare_(United_States)
6: 70235 United_States_district_court
7: 6097240 Minneapolis
8: 58353 Medicare_(United_States)
9: 58353 Medicare_(United_States)
10: 58353 Medicare_(United_States)
1: 120211 Minnetonka,_Minnesota
 [pem_pos=0, final_scores=0.46457154,
final_local_scores=0.5154584, pem=0.0, similarity=19.47202]
2: 19590 Minnesota
 [pem_pos=0, final_scores=0.721304,
final_local_scores=0.7596347, pem=0.0, similarity=26.758726]
3: 70235 United_States_district_court
 [pem_pos=0, final_scores=0.42257476,
final_local_scores=0.44837213, pem=0.0, similarity=17.470032]
4: 6097240 Minneapolis
 [pem_pos=0, final_scores=0.36348015,
final_local_scores=0.37900543, pem=-0.0128788,
similarity=15.470029]
5: 58353 Medicare_(United_States)
 [pem_pos=0, final_scores=0.5254462,
final_local_scores=0.5320875, pem=0.0, similarity=19.968266]
6: 70235 United_States_district_court
 [pem_pos=0, final_scores=0.47149783,
final_local_scores=0.48669708, pem=-0.01716152,

```
similarity=18.707052]
7: 6097240 Minneapolis
    [pem_pos=0, final_scores=0.43951774,
final_local_scores=0.46206492, pem=-0.0128788,
similarity=17.948689]
8: 58353 Medicare_(United_States)
    [pem_pos=0, final_scores=0.4032594,
final_local_scores=0.39861703, pem=0.0, similarity=15.98524]
9: 58353 Medicare_(United_States)
    [pem_pos=0, final_scores=0.45199418,
final_local_scores=0.45185238, pem=0.0, similarity=17.573889]
10: 58353 Medicare_(United_States)
    [pem_pos=0, final_scores=0.580032,
final_local_scores=0.5917142, pem=0.0, similarity=21.747644]
```

Sample: b'65_Oppenheimer'

Oppenheimer Capital to review Oct. div . [1 1 NEW YORK 1] 1996-08-22 [2 2 Oppenheimer Capital LP 2] said on Thursday it will review its cash distribution rate for the October quarterly distribution , assuming continued favorable results . The company , which reported improved first quarter earnings for the period ended July 31 , 1996 , declared a quarterly distribution of \$ 0.65 per partnership unit for the quarter ended July .

```
1: 645042 New_York_City
2: 16001392 Oppenheimer_Holdings
1: 645042 New_York_City
    [pem_pos=1, final_scores=0.32137173,
final_local_scores=0.40992433, pem=-1.7906687,
similarity=26.060608]
2: 16001392 Oppenheimer_Holdings
    [pem_pos=0, final_scores=0.38984352,
final_local_scores=0.4847194, pem=0.0, similarity=18.554707]
```

Sample: b'66_Best'

[1 1 Best 1] sees Q2 loss similar to Q1 loss . [2 2 RICHMOND 2] , Va . 1996-08-22 Best Products Co Chairman and Chief Executive Daniel Levy said Thursday he expected the company 's second-quarter results to be similar to the \$ 34.6 million loss posted in the first quarter . He also told [3 3 Reuters 3] before the retailer 's annual meeting that the second quarter could be better than the first quarter ended May 4 . " We could do even better , " he said . The second-quarter results are expected to be released in September . Levy said seeking bankruptcy protection was not under consideration . [4 4 Best 4] emerged from Chapter 11 bankruptcy protection in June 1994 after 3-1/2 years . " Bankruptcy is always possible , particularly when you lose what we are going to lose in the first half of this year , " Levy said . " But this is not something we are striving to do . " The Richmond-based retailer lost \$ 95.7 million in the fiscal year ended February 3 . That was the second-largest loss in the company

's history . Levy said that [5 5 Best 5] planned to open two new stores this fall . The company announced in March that it was closing seven stores and backing out of nine new lease agreements . At the time , [6 6 Best 6] said it did not plan to open any new stores this fall . It currently operates 169 stores in 23 states . For last year 's second quarter , which ended July 29 , 1995 , [7 7 Best 7] posted a loss of \$ 7.1 million , or \$ 0.23 per share , on sales of \$ 311.9 million .

1: 2691905 Best_Products
2: 53274 Richmond,_Virginia
3: 18998750 Reuters
4: 2691905 Best_Products
5: 2691905 Best_Products
6: 2691905 Best_Products
7: 2691905 Best_Products
1: 2691905 Best_Products
 [pem_pos=9, final_scores=0.81548995,
final_local_scores=0.78130573, pem=-3.7158177,
similarity=47.61263]
2: 53274 Richmond,_Virginia
 [pem_pos=1, final_scores=0.37118667,
final_local_scores=0.44586247, pem=-1.7896534,
similarity=27.127552]
3: 18998750 Reuters
 [pem_pos=0, final_scores=0.27291715,
final_local_scores=0.4218536, pem=0.0, similarity=16.678665]
4: 2691905 Best_Products
 [pem_pos=9, final_scores=0.816897,
final_local_scores=0.7828427, pem=-3.7158177,
similarity=47.658493]
5: 2691905 Best_Products
 [pem_pos=9, final_scores=0.9447971,
final_local_scores=0.9225542, pem=-3.7158177,
similarity=51.827763]
6: 2691905 Best_Products
 [pem_pos=9, final_scores=0.8839939,
final_local_scores=0.8561359, pem=-3.7158177,
similarity=49.84571]
7: 2691905 Best_Products
 [pem_pos=9, final_scores=0.88805145,
final_local_scores=0.8605681, pem=-3.7158177,
similarity=49.977974]

Sample: b'67_Measles'

Measles exposure can lead to bowel disease - study . [1 1 L0 NDON 1] 1996-08-23 Women who get measles while pregnant may have babies at higher risk of [2 2 Crohn 2] 's disease , a debilitating bowel disorder , researchers said on Friday . Three out of four [3 3 Swedish 3] babies born to mothers who caught measles developed serious cases of [4 4 Crohn 4] 's disease , the researchers said . Dr [5 5 Andrew Wakefield 5] of the [6 6 Royal Free Hospital School of Medicine 6] and

colleagues screened 25,000 babies delivered at [7 7 University Hospital 7] , [8 8 Uppsala 8] , between 1940 and 1949 . Four of the mothers had measles while pregnant . " Three of the four children had [9 9 Crohn 9] 's disease , " [10 10 Wakefield 10] 's group wrote in the [11 11 Lancet 11] medical journal . [12 12 Crohn 12] 's is an inflammation of the bowel that can sometimes require surgery . It causes diarrhoea , abdominal pain and weight loss . The researchers said the three children involved had especially severe cases . Exposure to viruses can often cause birth defects . Most notably , women who get rubella ([13 13 German 13] measles) have a high risk of a stillborn baby .

```
1: 17867 London
2: 63522 Crohn's_disease
3: 5058739 Sweden
4: 63522 Crohn's_disease
5: 19990546 Andrew_Wakefield
6: 1250182 Royal_Free_Hospital
7: 7946227 Uppsala_University_Hospital
8: 31784 Uppsala
9: 63522 Crohn's_disease
10: 19990546 Andrew_Wakefield
11: 463899 The_Lancet
12: 63522 Crohn's_disease
13: 11867 Germany
1: 17867 London
    [pem_pos=0, final_scores=0.65436405,
final_local_scores=0.7326136, pem=-0.09086737,
similarity=26.446514]
2: 63522 Crohn's_disease
    [pem_pos=0, final_scores=0.5418187,
final_local_scores=0.5574776, pem=0.0, similarity=20.725956]
3: 5058739 Sweden
    [pem_pos=0, final_scores=0.5667954,
final_local_scores=0.65025014, pem=-0.63135606,
similarity=26.927887]
4: 63522 Crohn's_disease
    [pem_pos=0, final_scores=0.5520856,
final_local_scores=0.5686926, pem=0.0, similarity=21.060635]
5: 19990546 Andrew_Wakefield
    [pem_pos=0, final_scores=0.5637682,
final_local_scores=0.589235, pem=0.0, similarity=21.67366]
6: 1250182 Royal_Free_Hospital
    [pem_pos=1, final_scores=0.356292,
final_local_scores=0.39249325, pem=-0.6325226,
similarity=19.24225]
7: 7946227 Uppsala_University_Hospital
    [pem_pos=5, final_scores=0.3993889,
final_local_scores=0.49987394, pem=-2.3576221,
similarity=31.828066]
8: 31784 Uppsala
    [pem_pos=0, final_scores=0.3412271,
```

```
final_local_scores=0.38467467, pem=-0.016858567,
similarity=15.660851]
9: 63522 Crohn's_disease
    [pem_pos=0, final_scores=0.62688,
final_local_scores=0.6503942, pem=0.0, similarity=23.498774]
10: 19990546 Andrew_Wakefield
    [pem_pos=0, final_scores=1.2072536,
final_local_scores=1.3129342, pem=0.0, similarity=43.270287]
11: 463899 The_Lancet
    [pem_pos=0, final_scores=0.35239524,
final_local_scores=0.3675179, pem=0.0, similarity=15.057179]
12: 63522 Crohn's_disease
    [pem_pos=0, final_scores=0.4157074,
final_local_scores=0.41972023, pem=0.0, similarity=16.615002]
13: 11867 Germany
    [pem_pos=1, final_scores=0.5969317,
final_local_scores=0.6861749, pem=-1.0465823,
similarity=30.258018]
```

Sample: b'68_All'

All the key numbers - [1 1 CBI 1] August industrial trends .
[2 2 LONDON 2] 1996-08-23 Following are key data from the
August monthly survey of trends in UK manufacturing by the [3
3 Confederation of British Industry 3] ([4 4 CBI 4]) . [5 5
CBI 5] MONTHLY TRENDS ENQUIRY (a) AUG JULY JUNE MAY - total
order book - 10 - 22 - 13 - 17 - export order book - 14 - 13
- 11 - 7 - stocks of finished goods +17 +19 +17 +25 - output
expectations * +22 +22 +12 +16 - domestic price expectations
* 0 - 1 +6 +4 NOTES - * over the coming four months ; - (a)
in percent , giving balance between those replying " above
normal " and those replying " below normal . " The survey was
conducted between July 23 and August 14 and involved 1,305
companies , representing 50 industries , accounting for
around half of the UK 's manufactured exports and some two
million employees . -- Rosemary Bennett , London Newsroom +44
171 542 7715

```
1: 1165269 Confederation_of_British_Industry
2: 17867 London
3: 1165269 Confederation_of_British_Industry
4: 1165269 Confederation_of_British_Industry
5: 1165269 Confederation_of_British_Industry
1: 1165269 Confederation_of_British_Industry
    [pem_pos=2, final_scores=0.53750503,
final_local_scores=0.46961147, pem=-1.2907379,
similarity=25.123089]
2: 17867 London
    [pem_pos=0, final_scores=0.6139823,
final_local_scores=0.72129244, pem=-0.09086737,
similarity=26.108667]
3: 1165269 Confederation_of_British_Industry
    [pem_pos=0, final_scores=0.6188392,
final_local_scores=0.55845666, pem=0.0, similarity=20.755173]
```

```
4: 1165269 Confederation_of_British_Industry
    [pem_pos=2, final_scores=0.56925184,
final_local_scores=0.50429, pem=-1.2907379,
similarity=26.157965]
5: 1165269 Confederation_of_British_Industry
    [pem_pos=2, final_scores=0.5431362,
final_local_scores=0.47576267, pem=-1.2907379,
similarity=25.306654]

Sample: b'69_London'
[1 1 London 1] shipsales . [2 2 LONDON 2] 1996-08-22 -
Secondhand tonnage brokers reported the sale of the following
vessels . Iron Gippsland - ( built 1989 ) 87,241 dwt sold to [3 3 Greek 3] buyers for $ 30 million . Sairyu Maru No : 2 -
( built 1982 ) 60,960 dwt sold to [4 4 Greek 4] buyers for $ 15.5 million . Stainless Fighter - ( built 1970 ) 21,718 dwt sold at auction for $ 6 million . Some of these sales may not be final as they may be subject to inspection , survey or other conditions .
1: 17867 London
2: 17867 London
3: 12108 Greece
4: 12108 Greece
1: 17867 London
    [pem_pos=0, final_scores=0.6780574,
final_local_scores=0.67091393, pem=-0.09086737,
similarity=24.605272]
2: 17867 London
    [pem_pos=0, final_scores=0.6859097,
final_local_scores=0.6794914, pem=-0.09086737,
similarity=24.861242]
3: 12108 Greece
    [pem_pos=1, final_scores=0.73935944,
final_local_scores=0.73787713, pem=-2.3247876,
similarity=38.751995]
4: 12108 Greece
    [pem_pos=1, final_scores=0.723023,
final_local_scores=0.720032, pem=-2.3247876,
similarity=38.21946]
```

```
Sample: b'70_Garlic'
Garlic pills do n't lower cholesterol , study finds . [1 1 L0
NDON 1] 1996-08-22 Garlic pills may not lower blood
cholesterol and studies that show they do may be flawed , [2 2
British 2] researchers have reported . A study by a team of
doctors at [3 3 Oxford University 3] has found people with
high blood cholesterol do not benefit significantly from
taking garlic tablets . The study involved 115 people with
high blood cholesterol levels . They were given 900
milligrams a day of dried garlic powder or placebo tablets .
" There were no significant differences between the groups
receiving garlic and placebo , " they wrote in the Journal of
```

the Royal College of Physicians . Those taking part were told to eat a low-fat diet for six weeks before they started taking the pills , and their blood cholesterol measured before and after the six-week period . The researchers said this would make their findings more accurate . Several studies have found garlic pills can lower blood pressure and blood cholesterol . But the [4 4 Oxford 4] team disputed these findings and said either previous trials may have been interpreted incorrectly , those taking part were not given special diets beforehand or the duration of the studies may have been too short . The six-month trial was funded by the [5 5 British Heart Foundation 5] and Lichtwer Pharma GmbH , which makes Kwai brand garlic tablets . The study did not address whether whole garlic could affect cholesterol . -- London Newsroom +44 171 542 7950

```
1: 17867 London
2: 31717 United_Kingdom
3: 31797 University_of_Oxford
4: 31797 University_of_Oxford
5: 1489428 British_Heart_Foundation
1: 17867 London
    [pem_pos=0, final_scores=0.6637934,
final_local_scores=0.66412365, pem=-0.09086737,
similarity=24.402637]
2: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5056831,
final_local_scores=0.5082741, pem=-0.48936236,
similarity=21.918854]
3: 31797 University_of_Oxford
    [pem_pos=0, final_scores=0.39582068,
final_local_scores=0.35572094, pem=-0.12160307,
similarity=15.366431]
4: 31797 University_of_Oxford
    [pem_pos=1, final_scores=0.309727,
final_local_scores=0.26167673, pem=-1.4798834,
similarity=19.946505]
5: 1489428 British_Heart_Foundation
    [pem_pos=0, final_scores=0.42420423,
final_local_scores=0.4916144, pem=0.0, similarity=18.760468]
```

Sample: b'71_Britain'

```
[1 1 Britain 1] gives aid to volcano-hit [2 2 Caribbean 2]
island . [3 3 LONDON 3] 1996-08-22 [4 4 Britain 4] said on
Thursday it would give 25 million pounds ( $ 39 million ) of
development aid to the [5 5 Caribbean 5] island of [6 6 Monts
errat 6] , where much of the population living in the south
has fled to avoid a volcano . The volcano in the Soufriere
hills has erupted three times in the past 13 months and last
April some 4,500 people living in the capital , [7 7 Plymouth
7] , and southern areas were evacuated to the north , where
many are living in public shelters and schools . " This
assistance will provide a fast track development programme
```

for the designated (northern) safe area , " [8 8 Britain 8] 's [9 9 Overseas Development Administration 9] said in a statement . [10 10 Britain 10] gave 8.5 million pounds (\$ 13 million) to [11 11 Montserrat 11] , which is one of its dependent territories , when the volcano first became active . Overseas Development Minister [12 12 Lynda Chalker 12] said a recent census had shown most Montserratians wanted to remain on the island . " The development of the north will help them to do just that , " she said .

1: 31717 United_Kingdom

2: 18956035 Caribbean

3: 17867 London

4: 31717 United_Kingdom

5: 18956035 Caribbean

6: 19281 Montserrat

7: 376904 Plymouth,_Montserrat

8: 31717 United_Kingdom

9: 247311 Department_for_International_Development

10: 31717 United_Kingdom

11: 19281 Montserrat

12: 2962836 Lynda_Chalker,_Baroness_Chalker_of_Wallasey

1: 31717 United_Kingdom

[pem_pos=0, final_scores=0.5762771,
final_local_scores=0.53198797, pem=-0.61114657,
similarity=23.288805]

2: 18956035 Caribbean

[pem_pos=0, final_scores=0.447514,
final_local_scores=0.43233776, pem=-0.09554641,
similarity=17.511131]

3: 17867 London

[pem_pos=0, final_scores=0.6402797,
final_local_scores=0.64228815, pem=-0.09086737,
similarity=23.751022]

4: 31717 United_Kingdom

[pem_pos=0, final_scores=0.5866024,
final_local_scores=0.54326683, pem=-0.61114657,
similarity=23.62539]

5: 18956035 Caribbean

[pem_pos=0, final_scores=0.37477273,
final_local_scores=0.35287905, pem=-0.09554641,
similarity=15.139925]

6: 19281 Montserrat

[pem_pos=0, final_scores=0.41824222,
final_local_scores=0.40264505, pem=-0.3910392,
similarity=18.231977]

7: 376904 Plymouth,_Montserrat

[pem_pos=12, final_scores=0.39845556,
final_local_scores=0.5059026, pem=-3.8796058,
similarity=40.284756]

8: 31717 United_Kingdom

[pem_pos=0, final_scores=0.50794023,
final_local_scores=0.4573403, pem=-0.61114657,

```
similarity=21.061169]
9: 247311 Department_for_International_Development
    [pem_pos=0, final_scores=0.42834634,
final_local_scores=0.4830767, pem=0.0, similarity=18.505686]
10: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5610538,
final_local_scores=0.51535887, pem=-0.61114657,
similarity=22.792559]
11: 19281 Montserrat
    [pem_pos=0, final_scores=0.40791422,
final_local_scores=0.39136326, pem=-0.3910392,
similarity=17.895306]
12: 2962836 Lynda_Chalker,_Baroness_Chalker_of_Wallasey
    [pem_pos=0, final_scores=0.4744051,
final_local_scores=0.596427, pem=0.0, similarity=21.888285]
```

Sample: b'72_Tennis'

Tennis - [1 1 Philippoussis 1] looms for [2 2 Sampras 2] in [3 3 U.S. Open 3] . Bill Berkrot [4 4 NEW YORK 4] 1996-08-22 World number one [5 5 Pete Sampras 5] , seeking his first [6 6 Grand Slam 6] title of the year , and women 's top seed [7 7 Steffi Graf 7] , aiming for her third , should be able to ease into the year 's final major , which begins on Monday . [8 8 Sampras 8] opens the defence of his [9 9 U.S. Open 9] crown against [10 10 David Rikl 10] of the [11 11 Czech Republic 11] , while top-ranked [12 12 Graf 12] begins her title defence against [13 13 Yayuk Basuki 13] of [14 14 Indonesia 14] . Wednesday 's [15 15 U.S. Open 15] draw ceremony revealed that both title holders should run into their first serious opposition in the third round . Looming in [16 16 Sampras 16] 's future is a likely third-round date with recent nemesis [17 17 Mark Philippoussis 17] , the rising [18 18 Australian 18] who took out [19 19 Sampras 19] in the third round of the [20 20 Australian Open 20] in January . [21 21 Sampras 21] avenged that defeat with a straight sets win over the 19-year-old power hitter in the second round at [22 22 Wimbledon 22] and their rubber match in [23 23 New York 23] could provide some first-week fireworks . While only a stunning upset will keep [24 24 Graf 24] from sailing through to a predictable semifinal showdown with third seed [25 25 Arantxa Sanchez Vicario 25] , the [26 26 German 26] star could also be tested in the third round where she will probably face 28th-ranked veteran [27 27 Natasha Zvereva 27] of [28 28 Belarus 28] . There will be no repeat of last year 's men 's final with eighth-ranked [29 29 Andre Agassi 29] landing in [30 30 Sampras 30] 's half of the draw . Bumping [31 31 Agassi 31] up to the sixth seeding avoided the possibility that he would run into [32 32 Sampras 32] as early as the quarter-finals , but they could lock horns in the semis . [33 33 Olympic 33] champion [34 34 Agassi 34] meets [35 35 Karim Alami 35] of [36 36 Morocco 36] in the first round . Surprise second seed [37 37 Michael Chang 37] , ranked third in the world , opens

against [38 38 Czech 38] [39 39 Daniel Vacek 39] , while women 's second seed [40 40 Monica Seles 40] drew [41 41 American 41] Anne Miller as her first victim . Second-ranked [42 42 Austrian 42] [43 43 Thomas Muster 43] , who was seeded third , did not have the luck of the draw with him . In the first round [44 44 Muster 44] faces [45 45 American 45] [46 46 Richey Reneberg 46] , who has been playing some of the best tennis of his career of late . If he survives , [47 47 Muster 47] is seeded to run into either fifth-seeded [48 48 Wimbledon 48] champion [49 49 Richard Krajicek 49] of the [50 50 Netherlands 50] or 12th-seeded [51 51 American 51] [52 52 Todd Martin 52] in the quarter-finals in [53 53 Chang 53] 's half of the draw . Perhaps the best , yet most unfortunate , first-round matchup of the men 's competition pits eighth seed [54 54 Jim Courier 54] against retiring star [55 55 Stefan Edberg 55] . The popular [56 56 Swede 56] is playing his final major tournament next week and the two-time champion 's [57 57 Grand Slam 57] farewell could well be a one-match affair . With the exception of a [58 58 Philippoussis 58] showdown , [59 59 Sampras 59] looks to have landed in a comfortable quarter of the draw with the likes of [60 60 Frenchman 60] [61 61 Cedric Pioline 61] and ailing [62 62 French Open 62] champion [63 63 Yevgeny Kafelnikov 63] , who is nursing a rib injury , in his path . [64 64 Seles 64] , runner-up to [65 65 Graf 65] last year , is seeded to run into fifth-ranked [66 66 German 66] [67 67 Anke Huber 67] in the quarter-finals with fourth seed [68 68 Conchita Martinez 68] or eighth-seeded [69 69 Olympic 69] champion [70 70 Lindsay Davenport 70] looking like her most likely semifinal opponents . But [71 71 Huber 71] will be tested immediately with a first-round encounter against dangerous 18th-ranked [72 72 South African 72] [73 73 Amanda Coetzer 73] . Sanchez Vicario , runner-up to [74 74 Graf 74] at the [75 75 French Open 75] and [76 76 Wimbledon 76] , begins play against a qualifier in a quarter of the draw that includes young talent [77 77 Martina Hingis 77] , the 16th seed , before a probable quarter-final clash with seventh-seeded veteran [78 78 Jana Novotna 78] . [79 79 Martinez 79] begins play against [80 80 Ruxandra Dragomir 80] of [81 81 Romania 81] . -- Sydney Newsroom 61-2 9373-1800

- 1: 261368 Mark_Phillipoussis
- 2: 65961 Pete_Sampras
- 3: 82729 US_Open_(tennis)
- 4: 645042 New_York_City
- 5: 65961 Pete_Sampras
- 6: 197638 Grand_Slam_(tennis)
- 7: 164640 Steffi_Graf
- 8: 65961 Pete_Sampras
- 9: 82729 US_Open_(tennis)
- 10: 10250905 David_Rikl
- 11: 5321 Czech_Republic
- 12: 164640 Steffi_Graf
- 13: 2672548 Yayuk_Basuki

14: 14579 Indonesia
15: 82729 US_Open_(tennis)
16: 65961 Pete_Sampras
17: 261368 Mark_Philippoussis
18: 4689264 Australia
19: 65961 Pete_Sampras
20: 97851 Australian_Open
21: 65961 Pete_Sampras
22: 91195 The_Championships,_Wimbledon
23: 645042 New_York_City
24: 164640 Steffi_Graf
25: 534624 Arantxa_Sánchez_Vicario
26: 11867 Germany
27: 916019 Natasha_Zvereva
28: 3457 Belarus
29: 595 Andre_Agassi
30: 65961 Pete_Sampras
31: 595 Andre_Agassi
32: 65961 Pete_Sampras
33: 27566 Summer_Olympic_Games
34: 595 Andre_Agassi
35: 3930942 Karim_Alami
36: 19291 Morocco
37: 391834 Michael_Chang
38: 5321 Czech_Republic
39: 3930665 Daniel_Vacek
40: 300468 Monica_Seles
41: 3434750 United_States
42: 26964606 Austria
43: 166170 Thomas_Muster
44: 166170 Thomas_Muster
45: 3434750 United_States
46: 871728 Richey_Reneberg
47: 166170 Thomas_Muster
48: 91195 The_Championships,_Wimbledon
49: 161972 Richard_Krajicek
50: 21148 Netherlands
51: 3434750 United_States
52: 871725 Todd_Martin
53: 391834 Michael_Chang
54: 685780 Jim_Courier
55: 391932 Stefan_Edberg
56: 5058739 Sweden
57: 197638 Grand_Slam_(tennis)
58: 261368 Mark_Philippoussis
59: 65961 Pete_Sampras
60: 5843419 France
61: 979076 Cédric_Pioline
62: 147724 French_Open
63: 489963 Yevgeny_Kafelnikov
64: 300468 Monica_Seles
65: 164640 Steffi_Graf

```
66: 11867 Germany
67: 1443478 Anke_Huber
68: 784152 Conchita_Martínez
69: 27566 Summer_Olympic_Games
70: 52828 Lindsay_Davenport
71: 1443478 Anke_Huber
72: 17416221 South_Africa
73: 1951841 Amanda_Coetzer
74: 164640 Steffi_Graf
75: 147724 French_Open
76: 91195 The_Championships,_Wimbledon
77: 19047 Martina_Hingis
78: 307050 Jana_Novotná
79: 784152 Conchita_Martínez
80: 4038121 Ruxandra_Dragomir
81: 25445 Romania
1: 261368 Mark_Philippoussis
    [pem_pos=0, final_scores=0.6602996,
final_local_scores=0.66332376, pem=0.0, similarity=23.884617]
2: 65961 Pete_Sampras
    [pem_pos=0, final_scores=0.6637355,
final_local_scores=0.63148403, pem=0.0, similarity=22.934456]
3: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.5277768,
final_local_scores=0.50297403, pem=-1.2370842,
similarity=25.826918]
4: 645042 New_York_City
    [pem_pos=1, final_scores=0.2989512,
final_local_scores=0.37433243, pem=-1.7906687,
similarity=24.998474]
5: 65961 Pete_Sampras
    [pem_pos=0, final_scores=0.7209056,
final_local_scores=0.6939337, pem=0.0, similarity=24.798079]
6: 197638 Grand_Slam_(tennis)
    [pem_pos=0, final_scores=0.5398957,
final_local_scores=0.51077855, pem=-0.37449613,
similarity=21.368933]
7: 164640 Steffi_Graf
    [pem_pos=0, final_scores=0.82537407,
final_local_scores=0.81868255, pem=0.0, similarity=28.520834]
8: 65961 Pete_Sampras
    [pem_pos=0, final_scores=0.73150635,
final_local_scores=0.7055134, pem=0.0, similarity=25.14364]
9: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.55637264,
final_local_scores=0.5342107, pem=-1.2370842,
similarity=26.75908]
10: 10250905 David_Rikl
    [pem_pos=0, final_scores=0.6848184,
final_local_scores=0.74133396, pem=0.0, similarity=26.212597]
11: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.5004557,
```

```
final_local_scores=0.5706959, pem=-0.27794555,  
similarity=22.631927]  
12: 164640 Steffi_Graf  
    [pem_pos=0, final_scores=1.2498323,  
final_local_scores=1.2835662, pem=0.0, similarity=42.39389]  
13: 2672548 Yayuk_Basuki  
    [pem_pos=0, final_scores=0.86607534,  
final_local_scores=0.93162626, pem=0.0, similarity=31.891298]  
14: 14579 Indonesia  
    [pem_pos=0, final_scores=0.59034747,  
final_local_scores=0.72459304, pem=-0.23151328,  
similarity=26.972021]  
15: 82729 US_Open_(tennis)  
    [pem_pos=1, final_scores=0.5532931,  
final_local_scores=0.5308468, pem=-1.2370842,  
similarity=26.658695]  
16: 65961 Pete_Sampras  
    [pem_pos=0, final_scores=0.63585466,  
final_local_scores=0.60102844, pem=0.0, similarity=22.0256]  
17: 261368 Mark_Philippoussis  
    [pem_pos=0, final_scores=0.7172286,  
final_local_scores=0.72551, pem=0.0, similarity=25.74038]  
18: 4689264 Australia  
    [pem_pos=0, final_scores=0.6007814,  
final_local_scores=0.7127624, pem=-0.491009,  
similarity=28.030148]  
19: 65961 Pete_Sampras  
    [pem_pos=0, final_scores=0.63297856,  
final_local_scores=0.59788674, pem=0.0, similarity=21.931847]  
20: 97851 Australian_Open  
    [pem_pos=0, final_scores=0.40953636,  
final_local_scores=0.37579268, pem=-0.3261775,  
similarity=17.077919]  
21: 65961 Pete_Sampras  
    [pem_pos=0, final_scores=0.6493755,  
final_local_scores=0.61579794, pem=0.0, similarity=22.466352]  
22: 91195 The_Championships,_Wimbledon  
    [pem_pos=0, final_scores=0.46497154,  
final_local_scores=0.4287867, pem=-0.8505555,  
similarity=21.511017]  
23: 645042 New_York_City  
    [pem_pos=1, final_scores=0.38278383,  
final_local_scores=0.46590686, pem=-1.7906687,  
similarity=27.731237]  
24: 164640 Steffi_Graf  
    [pem_pos=0, final_scores=1.3598952,  
final_local_scores=1.4037933, pem=0.0, similarity=45.981705]  
25: 534624 Arantxa_Sánchez_Vicario  
    [pem_pos=0, final_scores=0.49052292,  
final_local_scores=0.4801759, pem=0.0, similarity=18.41912]  
26: 11867 Germany  
    [pem_pos=1, final_scores=0.6575075,
```

```
final_local_scores=0.7502001, pem=-1.0465823,
similarity=32.16866]
27: 916019 Natasha_Zvereva
    [pem_pos=0, final_scores=0.8294764,
final_local_scores=0.87215394, pem=0.0, similarity=30.116526]
28: 3457 Belarus
    [pem_pos=0, final_scores=0.4630391,
final_local_scores=0.5443206, pem=-0.4220667,
similarity=22.62859]
29: 595 Andre_Agassi
    [pem_pos=0, final_scores=0.70375377,
final_local_scores=0.68103087, pem=0.0, similarity=24.413033]
30: 65961 Pete_Sampras
    [pem_pos=0, final_scores=0.6788293,
final_local_scores=0.64797175, pem=0.0, similarity=23.426483]
31: 595 Andre_Agassi
    [pem_pos=0, final_scores=0.68253046,
final_local_scores=0.65784764, pem=0.0, similarity=23.721199]
32: 65961 Pete_Sampras
    [pem_pos=0, final_scores=0.6571353,
final_local_scores=0.6242743, pem=0.0, similarity=22.719303]
33: 27566 Summer_Olympic_Games
    [pem_pos=2, final_scores=0.38213354,
final_local_scores=0.41541713, pem=-3.2532303,
similarity=34.17816]
34: 595 Andre_Agassi
    [pem_pos=0, final_scores=0.61611015,
final_local_scores=0.5852935, pem=0.0, similarity=21.556038]
35: 3930942 Karim_Alami
    [pem_pos=0, final_scores=0.6220856,
final_local_scores=0.7110967, pem=0.0, similarity=25.310257]
36: 19291 Morocco
    [pem_pos=0, final_scores=0.5340175,
final_local_scores=0.65242726, pem=-0.29598162,
similarity=25.169037]
37: 391834 Michael_Chang
    [pem_pos=0, final_scores=0.8748291,
final_local_scores=0.88064986, pem=0.0, similarity=30.370062]
38: 5321 Czech_Republic
    [pem_pos=1, final_scores=0.58265465,
final_local_scores=0.66048574, pem=-1.208016,
similarity=30.369303]
39: 3930665 Daniel_Vacek
    [pem_pos=0, final_scores=0.76747185,
final_local_scores=0.84489363, pem=0.0, similarity=29.303024]
40: 300468 Monica_Seles
    [pem_pos=0, final_scores=0.75283617,
final_local_scores=0.7456971, pem=0.0, similarity=26.342802]
41: 3434750 United_States
    [pem_pos=0, final_scores=0.52100736,
final_local_scores=0.61211276, pem=-0.12994486,
similarity=23.06304]
```

42: 26964606 Austria
[pem_pos=0, final_scores=0.66678715,
final_local_scores=0.7558324, pem=-0.40674555,
similarity=28.857204]
43: 166170 Thomas_Muster
[pem_pos=0, final_scores=0.9104312,
final_local_scores=0.92587465, pem=0.0, similarity=31.71966]
44: 166170 Thomas_Muster
[pem_pos=0, final_scores=1.0460875,
final_local_scores=1.0756416, pem=0.0, similarity=36.189007]
45: 3434750 United_States
[pem_pos=0, final_scores=0.5030307,
final_local_scores=0.592476, pem=-0.12994486,
similarity=22.47704]
46: 871728 Richey_Reneberg
[pem_pos=0, final_scores=0.72485334,
final_local_scores=0.78336126, pem=0.0, similarity=27.466776]
47: 166170 Thomas_Muster
[pem_pos=0, final_scores=1.0489818,
final_local_scores=1.0788033, pem=0.0, similarity=36.283356]
48: 91195 The_Championships,_Wimbledon
[pem_pos=0, final_scores=0.5598525,
final_local_scores=0.5324297, pem=-0.8505555,
similarity=24.60393]
49: 161972 Richard_Krajicek
[pem_pos=0, final_scores=0.89609796,
final_local_scores=0.93217, pem=0.0, similarity=31.907526]
50: 21148 Netherlands
[pem_pos=0, final_scores=0.5084982,
final_local_scores=0.5841593, pem=-0.24940795,
similarity=22.878511]
51: 3434750 United_States
[pem_pos=0, final_scores=0.63767534,
final_local_scores=0.7395549, pem=-0.12994486,
similarity=26.866167]
52: 871725 Todd_Martin
[pem_pos=0, final_scores=0.91895276,
final_local_scores=0.9450099, pem=0.0, similarity=32.29069]
53: 391834 Michael_Chang
[pem_pos=0, final_scores=1.069803,
final_local_scores=1.0950565, pem=0.0, similarity=36.768383]
54: 685780 Jim_Courier
[pem_pos=0, final_scores=0.8185578,
final_local_scores=0.8215863, pem=0.0, similarity=28.607487]
55: 391932 Stefan_Edberg
[pem_pos=0, final_scores=0.82472533,
final_local_scores=0.83090204, pem=0.0, similarity=28.885487]
56: 5058739 Sweden
[pem_pos=0, final_scores=0.426192,
final_local_scores=0.50183743, pem=-0.6464619,
similarity=22.581104]
57: 197638 Grand_Slam_(tennis)

[pem_pos=0, final_scores=0.5032056,
final_local_scores=0.4707002, pem=-0.37449613,
similarity=20.172915]
58: 261368 Mark_Philippoussis
[pem_pos=0, final_scores=0.6569833,
final_local_scores=0.65970117, pem=0.0, similarity=23.776512]
59: 65961 Pete_Sampras
[pem_pos=0, final_scores=0.673199,
final_local_scores=0.6418215, pem=0.0, similarity=23.242947]
60: 5843419 France
[pem_pos=1, final_scores=0.396833,
final_local_scores=0.4593575, pem=-0.7184992,
similarity=21.705168]
61: 979076 Cédric_Pioline
[pem_pos=0, final_scores=0.7000089,
final_local_scores=0.7223913, pem=0.0, similarity=25.64731]
62: 147724 French_Open
[pem_pos=0, final_scores=0.50728935,
final_local_scores=0.47693086, pem=-0.23693618,
similarity=19.610777]
63: 489963 Yevgeny_Kafelnikov
[pem_pos=0, final_scores=0.64544696,
final_local_scores=0.6531717, pem=0.0, similarity=23.58166]
64: 300468 Monica_Seles
[pem_pos=0, final_scores=0.72640496,
final_local_scores=0.71821004, pem=0.0, similarity=25.522533]
65: 164640 Steffi_Graf
[pem_pos=0, final_scores=1.2588153,
final_local_scores=1.2933788, pem=0.0, similarity=42.686714]
66: 11867 Germany
[pem_pos=1, final_scores=0.8041615,
final_local_scores=0.91039723, pem=-1.0465823,
similarity=36.94926]
67: 1443478 Anke_Huber
[pem_pos=0, final_scores=0.9282904,
final_local_scores=0.9814759, pem=0.0, similarity=33.37891]
68: 784152 Conchita_Martínez
[pem_pos=0, final_scores=0.8117338,
final_local_scores=0.84416354, pem=0.0, similarity=29.281237]
69: 27566 Summer_Olympic_Games
[pem_pos=2, final_scores=0.34139663,
final_local_scores=0.3709182, pem=-3.2532303,
similarity=32.850224]
70: 52828 Lindsay_Davenport
[pem_pos=0, final_scores=0.9363261,
final_local_scores=0.95006317, pem=0.0, similarity=32.441494]
71: 1443478 Anke_Huber
[pem_pos=0, final_scores=1.0090342,
final_local_scores=1.0720887, pem=0.0, similarity=36.082977]
72: 17416221 South_Africa
[pem_pos=0, final_scores=0.5980862,
final_local_scores=0.7325265, pem=-0.1389451,

```
similarity=26.705368]
73: 1951841 Amanda_Coetzer
    [pem_pos=0, final_scores=0.7519235,
final_local_scores=0.78252286, pem=0.0, similarity=27.441755]
74: 164640 Steffi_Graf
    [pem_pos=0, final_scores=1.2346573,
final_local_scores=1.26699, pem=0.0, similarity=41.899223]
75: 147724 French_Open
    [pem_pos=0, final_scores=0.46279252,
final_local_scores=0.42832482, pem=-0.23693618,
similarity=18.160275]
76: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.5567104,
final_local_scores=0.5289975, pem=-0.8505555,
similarity=24.501507]
77: 19047 Martina_Hingis
    [pem_pos=0, final_scores=0.71283674,
final_local_scores=0.70717895, pem=0.0, similarity=25.193344]
78: 307050 Jana_Novotná
    [pem_pos=0, final_scores=0.6986739,
final_local_scores=0.7149823, pem=0.0, similarity=25.42621]
79: 784152 Conchita_Martínez
    [pem_pos=0, final_scores=1.3659583,
final_local_scores=1.4517624, pem=0.0, similarity=47.413197]
80: 4038121 Ruxandra_Dragomir
    [pem_pos=0, final_scores=0.8257049,
final_local_scores=0.91291517, pem=0.0, similarity=31.332922]
81: 25445 Romania
    [pem_pos=0, final_scores=0.6647273,
final_local_scores=0.76256406, pem=-0.21241479,
similarity=28.00129]
```

Sample: b'73_RTRS'

RTRS - Tennis - [1 1 Muster 1] upset , [2 2 Philippoussis 2] wins , [3 3 Stoltenberg 3] loses . [4 4 TORONTO 4] 1996-08-21 Top-seeded [5 5 Thomas Muster 5] of [6 6 Austria 6] was beaten 6-3 7-5 by 123rd-ranked [7 7 Daniel Nestor 7] of [8 8 Canada 8] on Wednesday in his first match of the \$ 2 million [9 9 Canadian Open 9] . A lefthander with a strong serve , [10 10 Nestor 10] kept the rallies short by constantly attacking the net and the tactic worked in the second-round match against [1 11 Muster 11] , playing his first match after receiving a first-round bye along with the other top eight seeds . The tournament also lost its second seed on the third day of play when second-seeded [12 12 Goran Ivanisevic 12] of [13 13 Croatia 13] was beaten 6-7(3-7) 6-4 6-4 by unseeded [14 14 Mikael Tillstrom 14] of [15 15 Sweden 15] . Other seeded players advancing were number three [16 16 Wayne Ferreira 16] of [17 17 South Africa 17] , number four [18 18 Marcelo Rios 18] of [19 19 Chile 19] , number six [20 20 MaliVai Washington 20] of the [21 21 United States 21] and [22 22 American 22] [23 23 Todd Martin 23] , the seventh seed . Eighth seed [24 24 Marc Ro

sset 24] of [25 25 Switzerland 25] was eliminated in a one hour , 55 minute battle by unseeded [26 26 Mark Philippoussis 26] of [27 27 Australia 27] . [28 28 Philippoussis 28] saved a match point at 5-6 in the third-set tie break before winning 6-3 3-6 7-6 (8-6) . [29 29 Philippoussis 29] 's compatriot , 13th seed [30 30 Jason Stoltenberg 30] , was not as fortunate . He held one match point at 9-8 in a marathon third-set tie break but was beaten 5-7 7-6 (7-1) 7-6 (13-11) by unseeded [31 31 Daniel Vacek 31] of the [32 32 Czech Republic 32] . " I knew I had to serve well and keep the points short and that 's what I was able to do , " said [33 33 Nestor 33] , who ranks 10th in doubles . There were only two service breaks in the match . The lanky [34 34 Canadian 34] broke [35 35 Muster 35] at 4-3 in the first set and 5-5 in the second before ending the match on his third match point when the [36 36 Austrian 36] hit a service return long . " I probably did n't hit five ground strokes in the whole match , " said [37 37 Muster 37] , only partly joking . " The way he was chipping and charging and serving and volleying I did n't really get my timing playing from the baseline . " " He played a good match , took a few chances , and every time he was down he was able to come up with a big first serve . " Playing at night was not [38 38 Muster 38] 's preference . " I asked for a day match and they gave me a night match , " he said . " I do n't like playing under the lights but maybe it would not have made any difference . " [39 39 Ivanisevic 39] rallied from a 2-5 deficit in the first set but then played erratically against the 44th-ranked Tillstrom , who was a surprise winner over his famous compatriot [40 40 Stefan Edberg 40] in the second round at [41 41 Wimbledon 41] . [42 42 Ivanisevic 42] hit 32 aces but was outplayed from the back court by the 24-year-old Tillstrom . The sixth-ranked [43 43 Ivanisevic 43] , who lost in the final at [44 44 Indianapolis 44] to world number one [45 45 Pete Sampras 45] of the [46 46 U.S. 46] last Sunday , made a quick getaway after his loss but did say : " Something was not there when I arrived (in [47 47 Toronto 47]) . I did n't feel good . And I did n't have a good feeling as soon as I lost in my doubles (on Tuesday) . " " I thought he looked a little unfocused at certain times on his ground strokes , " said Tillstrom . The 19-year-old [48 48 Philippoussis 48] , who beat [49 49 Sampras 49] in the third round of this year 's [50 50 Australian Open 50] , stayed calm in a nervy third-set tie break against [51 51 Rosset 51] . " I 'm pleased because I did n't play that great today , but I fought really well , " he said . " When I was down 2-5 in the tiebreak (in the third set) , I just thought about winning my two serves and hoped that he might get tight . Then he shanked a forehand at to make it 5-all and that helped me back . "

1: 166170 Thomas_Muster
2: 261368 Mark_Philippoussis
3: 2598440 Jason_Stoltenberg

4: 64646 Toronto
5: 166170 Thomas_Muster
6: 26964606 Austria
7: 998494 Daniel_Nestor
8: 5042916 Canada
9: 1872219 Canadian_Open_(tennis)
10: 998494 Daniel_Nestor
11: 166170 Thomas_Muster
12: 194244 Goran_Ivanišević
13: 5573 Croatia
14: 3943969 Mikael_Tillström
15: 5058739 Sweden
16: 1723516 Wayne_Ferreira
17: 17416221 South_Africa
18: 483554 Marcelo_Ríos
19: 5489 Chile
20: 873616 MaliVai_Washington
21: 3434750 United_States
22: 3434750 United_States
23: 871725 Todd_Martin
24: 1717291 Marc_Rosset
25: 26748 Switzerland
26: 261368 Mark_Philippoussis
27: 4689264 Australia
28: 261368 Mark_Philippoussis
29: 261368 Mark_Philippoussis
30: 2598440 Jason_Stoltenberg
31: 3930665 Daniel_Vacek
32: 5321 Czech_Republic
33: 998494 Daniel_Nestor
34: 5042916 Canada
35: 166170 Thomas_Muster
36: 26964606 Austria
37: 166170 Thomas_Muster
38: 166170 Thomas_Muster
39: 194244 Goran_Ivanišević
40: 391932 Stefan_Edberg
41: 91195 The_Championships,_Wimbledon
42: 194244 Goran_Ivanišević
43: 194244 Goran_Ivanišević
44: 57707 Indianapolis
45: 65961 Pete_Sampras
46: 3434750 United_States
47: 64646 Toronto
48: 261368 Mark_Philippoussis
49: 65961 Pete_Sampras
50: 97851 Australian_Open
51: 1717291 Marc_Rosset
1: 166170 Thomas_Muster
 [pem_pos=0, final_scores=0.7882952,
final_local_scores=0.78405946, pem=0.0, similarity=27.487612]
2: 261368 Mark_Philippoussis

```
[pem_pos=0, final_scores=0.60677236,
final_local_scores=0.6011505, pem=0.0, similarity=22.029243]
3: 2598440 Jason_Stoltenberg
    [pem_pos=0, final_scores=0.72990245,
final_local_scores=0.76602346, pem=0.0, similarity=26.94938]
4: 64646 Toronto
    [pem_pos=0, final_scores=0.53373164,
final_local_scores=0.6210739, pem=-0.12291032,
similarity=23.292204]
5: 166170 Thomas_Muster
    [pem_pos=0, final_scores=1.0445421,
final_local_scores=1.061841, pem=0.0, similarity=35.777164]
6: 26964606 Austria
    [pem_pos=0, final_scores=0.6043749,
final_local_scores=0.6803702, pem=-0.25262418,
similarity=25.767126]
7: 998494 Daniel_Nestor
    [pem_pos=0, final_scores=0.82283634,
final_local_scores=0.8469094, pem=0.0, similarity=29.363178]
8: 5042916 Canada
    [pem_pos=0, final_scores=0.7063007,
final_local_scores=0.8108924, pem=-0.11662793,
similarity=28.9226]
9: 1872219 Canadian_Open_(tennis)
    [pem_pos=0, final_scores=0.42247158,
final_local_scores=0.41952085, pem=-0.58982956,
similarity=19.816637]
10: 998494 Daniel_Nestor
    [pem_pos=0, final_scores=1.0271477,
final_local_scores=1.0732713, pem=0.0, similarity=36.11827]
11: 166170 Thomas_Muster
    [pem_pos=0, final_scores=1.0320402,
final_local_scores=1.050314, pem=0.0, similarity=35.433178]
12: 194244 Goran_Ivanišević
    [pem_pos=0, final_scores=1.0433917,
final_local_scores=1.0708196, pem=0.0, similarity=36.045105]
13: 5573 Croatia
    [pem_pos=0, final_scores=0.5011386,
final_local_scores=0.5689113, pem=-0.35021794,
similarity=22.9717]
14: 3943969 Mikael_Tillström
    [pem_pos=0, final_scores=0.69013065,
final_local_scores=0.75869167, pem=0.0, similarity=26.730585]
15: 5058739 Sweden
    [pem_pos=0, final_scores=0.6870406,
final_local_scores=0.77226704, pem=-0.2135508,
similarity=28.297024]
16: 1723516 Wayne_Ferreira
    [pem_pos=0, final_scores=0.861409,
final_local_scores=0.9045438, pem=0.0, similarity=31.083105]
17: 17416221 South_Africa
    [pem_pos=0, final_scores=0.5563147,
```

```
final_local_scores=0.6770312, pem=-0.26251382,  
similarity=25.721266]  
18: 483554 Marcelo_Ríos  
    [pem_pos=0, final_scores=0.7697373,  
final_local_scores=0.78151375, pem=0.0, similarity=27.411642]  
19: 5489 Chile  
    [pem_pos=0, final_scores=0.49411213,  
final_local_scores=0.5937743, pem=-0.27404457,  
similarity=23.29942]  
20: 873616 MaliVai_Washington  
    [pem_pos=0, final_scores=0.7879315,  
final_local_scores=0.8328687, pem=0.0, similarity=28.944176]  
21: 3434750 United_States  
    [pem_pos=0, final_scores=0.4376161,  
final_local_scores=0.5138284, pem=-0.073457584,  
similarity=19.822851]  
22: 3434750 United_States  
    [pem_pos=0, final_scores=0.4652291,  
final_local_scores=0.54399145, pem=-0.12994486,  
similarity=21.030165]  
23: 871725 Todd_Martin  
    [pem_pos=0, final_scores=0.82916874,  
final_local_scores=0.8451517, pem=0.0, similarity=29.310726]  
24: 1717291 Marc_Rosset  
    [pem_pos=0, final_scores=0.78799516,  
final_local_scores=0.8323638, pem=0.0, similarity=28.929108]  
25: 26748 Switzerland  
    [pem_pos=0, final_scores=0.63803583,  
final_local_scores=0.7087054, pem=-0.18598008,  
similarity=26.250284]  
26: 261368 Mark_Philippoussis  
    [pem_pos=0, final_scores=0.8181059,  
final_local_scores=0.8320003, pem=0.0, similarity=28.918262]  
27: 4689264 Australia  
    [pem_pos=0, final_scores=0.6568761,  
final_local_scores=0.76416653, pem=-0.17168477,  
similarity=27.827616]  
28: 261368 Mark_Philippoussis  
    [pem_pos=0, final_scores=0.59518534,  
final_local_scores=0.58849347, pem=0.0, similarity=21.651531]  
29: 261368 Mark_Philippoussis  
    [pem_pos=0, final_scores=0.6651076,  
final_local_scores=0.66487294, pem=0.0, similarity=23.930847]  
30: 2598440 Jason_Stoltenberg  
    [pem_pos=0, final_scores=0.7772265,  
final_local_scores=0.8139308, pem=0.0, similarity=28.379032]  
31: 3930665 Daniel_Vacek  
    [pem_pos=0, final_scores=0.75874656,  
final_local_scores=0.83054703, pem=0.0, similarity=28.874893]  
32: 5321 Czech_Republic  
    [pem_pos=0, final_scores=0.49365544,  
final_local_scores=0.55593234, pem=-0.27794555,
```

```
similarity=22.191355]
33: 998494 Daniel_Nestor
    [pem_pos=0, final_scores=1.2781453,
final_local_scores=1.3474481, pem=0.0, similarity=44.30025]
34: 5042916 Canada
    [pem_pos=0, final_scores=0.40799314,
final_local_scores=0.48503667, pem=-0.41303828,
similarity=20.81034]
35: 166170 Thomas_Muster
    [pem_pos=0, final_scores=0.9409073,
final_local_scores=0.95076495, pem=0.0, similarity=32.462437]
36: 26964606 Austria
    [pem_pos=0, final_scores=0.42836958,
final_local_scores=0.48811114, pem=-0.40674555,
similarity=20.867868]
37: 166170 Thomas_Muster
    [pem_pos=0, final_scores=1.0390463,
final_local_scores=1.057967, pem=0.0, similarity=35.66156]
38: 166170 Thomas_Muster
    [pem_pos=0, final_scores=1.1341068,
final_local_scores=1.1618061, pem=0.0, similarity=38.760326]
39: 194244 Goran_Ivanišević
    [pem_pos=0, final_scores=0.77239233,
final_local_scores=0.7747939, pem=0.0, similarity=27.21111]
40: 391932 Stefan_Edberg
    [pem_pos=0, final_scores=0.83182293,
final_local_scores=0.83709437, pem=0.0, similarity=29.070278]
41: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.48074478,
final_local_scores=0.45996702, pem=-0.8505555,
similarity=22.4415]
42: 194244 Goran_Ivanišević
    [pem_pos=0, final_scores=0.56739825,
final_local_scores=0.550869, pem=0.0, similarity=20.528742]
43: 194244 Goran_Ivanišević
    [pem_pos=0, final_scores=0.56524855,
final_local_scores=0.54852086, pem=0.0, similarity=20.45867]
44: 57707 Indianapolis
    [pem_pos=0, final_scores=0.36514145,
final_local_scores=0.4336918, pem=-0.30184713,
similarity=18.673433]
45: 65961 Pete_Sampras
    [pem_pos=0, final_scores=0.6861103,
final_local_scores=0.66812694, pem=0.0, similarity=24.027954]
46: 3434750 United_States
    [pem_pos=0, final_scores=0.4683661,
final_local_scores=0.54741806, pem=-0.14075686,
similarity=21.191217]
47: 64646 Toronto
    [pem_pos=0, final_scores=0.5274859,
final_local_scores=0.6142514, pem=-0.12291032,
similarity=23.088606]
```

```

48: 261368 Mark_Philippoussis
    [pem_pos=0, final_scores=0.5852191,
final_local_scores=0.57760686, pem=0.0, similarity=21.326654]
49: 65961 Pete_Sampras
    [pem_pos=0, final_scores=0.6537789,
final_local_scores=0.6328098, pem=0.0, similarity=22.97402]
50: 97851 Australian_Open
    [pem_pos=0, final_scores=0.4204561,
final_local_scores=0.40306747, pem=-0.3261775,
similarity=17.891853]
51: 1717291 Marc_Rosset
    [pem_pos=0, final_scores=0.5826978,
final_local_scores=0.60810757, pem=0.0, similarity=22.236856]

```

Sample: b'74_Soccer'

Soccer - Results of [1 1 South Korean 1] pro-soccer games . [
2 2 SEOUL 2] 1996-08-22 Results of [3 3 South Korean 3] pro-
soccer games played on Wednesday . Anyang 3 Chonnam 3
(halftime 2-0) Puchon 0 [4 4 Suwon 4] 0 (halftime 0-0)
Standings after games played on Wednesday (tabulate under -
won , drawn , lost , goals for , goals against , points) : W
D L G / F G / A P Puchon 1 1 0 1 0 4 Chonan 1 0 0 5 4 3
Anyang 0 2 0 5 5 2 [5 5 Suwon 5] 0 2 0 3 3 2 [6 6 Pohang 6] 0
1 0 3 3 1 Pusan 0 1 0 0 2 1 Chonnam 0 1 1 5 6 1 [7 7 Ulsan 7]
0 0 1 4 5 0 Chonbuk 0 0 0 0 0 0

```

1: 27019 South_Korea
2: 19159283 Seoul
3: 27019 South_Korea
4: 1054817 Suwon_Samsung_Bluewings
5: 1054817 Suwon_Samsung_Bluewings
6: 1056366 Pohang_Steelers
7: 1056425 Ulsan_Hyundai_FC
1: 27019 South_Korea
    [pem_pos=0, final_scores=0.50410503,
final_local_scores=0.4979146, pem=0.0, similarity=18.948479]
2: 19159283 Seoul
    [pem_pos=0, final_scores=0.4471864,
final_local_scores=0.45521945, pem=-0.0378014,
similarity=17.87994]
3: 27019 South_Korea
    [pem_pos=0, final_scores=0.53057307,
final_local_scores=0.526827, pem=0.0, similarity=19.811281]
4: 1054817 Suwon_Samsung_Bluewings
    [pem_pos=2, final_scores=0.46060967,
final_local_scores=0.42967457, pem=-1.0623775,
similarity=22.689434]
5: 1054817 Suwon_Samsung_Bluewings
    [pem_pos=2, final_scores=0.51184326,
final_local_scores=0.4856395, pem=-1.0623775,
similarity=24.359537]
6: 1056366 Pohang_Steelers
    [pem_pos=2, final_scores=0.62057763,
```

```
final_local_scores=0.6301669, pem=-1.0587178,
similarity=28.652622]
7: 1056425 Ulsan_Hyundai_FC
    [pem_pos=3, final_scores=0.49684078,
final_local_scores=0.484711, pem=-1.7336445,
similarity=27.982285]

Sample: b'75_Senegal'
[1 1 Senegal 1] cholera outbreak kills five . [2 2 DAKAR 2]
1996-08-22 An outbreak of cholera has killed five people in
the central [3 3 Senegal 3] town of [4 4 Kaolack 4] , where
health authorities have recorded 291 cases since August 11 ,
a medical official said on Thursday . Doctor Masserigne
Ndiaye said medical staff were overwhelmed with work . "
People are rushing to the hospital as soon as the first
symptoms appear , that 's why we have fewer deaths , " he
told [5 5 Reuters 5] by telephone from the town , 160 km
( 100 miles ) southeast of the [6 6 Senegalese 6] capital [7 7
Dakar 7] .
1: 27278 Senegal
2: 44251 Dakar
3: 27278 Senegal
4: 1606283 Kaolack
5: 18998750 Reuters
6: 27278 Senegal
7: 44251 Dakar
1: 27278 Senegal
    [pem_pos=0, final_scores=0.6067505,
final_local_scores=0.56853485, pem=-0.40308514,
similarity=23.247967]
2: 44251 Dakar
    [pem_pos=0, final_scores=0.45142025,
final_local_scores=0.405883, pem=-0.008440792,
similarity=16.247972]
3: 27278 Senegal
    [pem_pos=0, final_scores=0.5233934,
final_local_scores=0.47747988, pem=-0.40308514,
similarity=20.530704]
4: 1606283 Kaolack
    [pem_pos=0, final_scores=0.40589088,
final_local_scores=0.45230222, pem=0.0, similarity=17.587313]
5: 18998750 Reuters
    [pem_pos=0, final_scores=0.32496256,
final_local_scores=0.41457027, pem=0.0, similarity=16.461317]
6: 27278 Senegal
    [pem_pos=0, final_scores=0.478669,
final_local_scores=0.42862523, pem=0.0, similarity=16.880745]
7: 44251 Dakar
    [pem_pos=0, final_scores=0.44679064,
final_local_scores=0.40082592, pem=-0.008440792,
similarity=16.09706]
```

Sample: b'76_Nigerian'

[1 1 Nigerian 1] general takes over [2 2 Liberia 2] [3 3 ECO
MOG 3] force . [4 4 MONROVIA 4] 1996-08-22 [5 5 Nigerian 5]
Major General Sam Victor Malu took over on Thursday as
commander of the [6 6 ECOMOG 6] peacekeeping force in [7 7 Li
beria 7] , two days after the start of the latest ceasefire
in the six-year civil war . Malu replaced another [8 8 Nigeri
an 8] major general , [9 9 John Inienger 9] , who told
officers at the handover ceremony that peace was now at hand
for [10 10 Liberia 10] after six years of fighting and more
than a dozen failed accords . " The search for peace in [11 11
Liberia 11] has been difficult , challenging and sometimes
painful . Peacekeepers were harassed , killed and taken
hostage , " he said . " It is difficult but I want to assure
you that peace is in sight . " [12 12 United Nations 12]
military observers travelling to the western town of [13 13 Tu
bmanburg 13] on Wednesday to monitor the ceasefire were
delayed by shooting along the highway , [14 14 U.N. 14]
special representative Anthony Nyakyi said . They finally
went ahead with an escort from the [15 15 ULIMO-J 15]
faction . Faction leaders who agreed a new peace deal in the [
16 16 Nigerian 16] capital [17 17 Abuja 17] on Saturday have
accused each other of breaking the ceasefire . The latest
peace deal foresees the disarmament of an estimated 60,000
combatants and sets a target date of May 30 next year for
elections . The [18 18 ECOMOG 18] force , currently 10,000
strong , was sent to [19 19 Liberia 19] by the [20 20 Economic
Community of West African States 20] in 1990 at the height of
the fighting .

1: 21383 Nigeria

2: 17791 Liberia

3: 1618193

Economic_Community_of_West_African_States_Monitoring_Group

4: 285213 Monrovia

5: 21383 Nigeria

6: 1618193

Economic_Community_of_West_African_States_Monitoring_Group

7: 17791 Liberia

8: 21383 Nigeria

9: 25574898 John_Mark_Inienger

10: 17791 Liberia

11: 17791 Liberia

12: 31769 United_Nations

13: 1561416 Tubmanburg

14: 31769 United_Nations

15: 2534235

United_Liberation_Movement_of_Liberia_for_Democracy_-
_Johnson_faction

16: 21383 Nigeria

17: 56823 Abuja

18: 1618193

Economic_Community_of_West_African_States_Monitoring_Group

```
19: 17791 Liberia
20: 36848 Economic_Community_of_West_African_States
1: 21383 Nigeria
    [pem_pos=0, final_scores=0.5621483,
final_local_scores=0.53962576, pem=-0.075027294,
similarity=20.601233]
2: 17791 Liberia
    [pem_pos=0, final_scores=0.57896084,
final_local_scores=0.534661, pem=-0.26536247,
similarity=21.488146]
3: 1618193
Economic_Community_of_West_African_States_Monitoring_Group
    [pem_pos=0, final_scores=0.5835328,
final_local_scores=0.5997972, pem=0.0, similarity=21.988857]
4: 285213 Monrovia
    [pem_pos=0, final_scores=0.42471927,
final_local_scores=0.41111404, pem=-0.2167236,
similarity=17.536753]
5: 21383 Nigeria
    [pem_pos=0, final_scores=0.50311756,
final_local_scores=0.47514373, pem=-0.075027294,
similarity=18.67696]
6: 1618193
Economic_Community_of_West_African_States_Monitoring_Group
    [pem_pos=0, final_scores=0.58698046,
final_local_scores=0.6035632, pem=0.0, similarity=22.101242]
7: 17791 Liberia
    [pem_pos=0, final_scores=0.50620323,
final_local_scores=0.45518434, pem=-0.26536247,
similarity=19.116404]
8: 21383 Nigeria
    [pem_pos=0, final_scores=0.5168207,
final_local_scores=0.49011236, pem=-0.075027294,
similarity=19.123653]
9: 25574898 John_Mark_Inienger
    [pem_pos=0, final_scores=0.5186854,
final_local_scores=0.65078413, pem=0.0, similarity=23.51041]
10: 17791 Liberia
    [pem_pos=0, final_scores=0.49183667,
final_local_scores=0.43949103, pem=-0.26536247,
similarity=18.648085]
11: 17791 Liberia
    [pem_pos=0, final_scores=0.5369577,
final_local_scores=0.4887789, pem=-0.26536247,
similarity=20.118933]
12: 31769 United_Nations
    [pem_pos=0, final_scores=0.42489582,
final_local_scores=0.43322605, pem=0.0, similarity=17.018042]
13: 1561416 Tubmanburg
    [pem_pos=0, final_scores=0.37743658,
final_local_scores=0.48860997, pem=0.0, similarity=18.670809]
14: 31769 United_Nations
```

```
[pem_pos=0, final_scores=0.51237017,
final_local_scores=0.52877855, pem=0.0, similarity=19.86952]
15: 2534235
United_Liberation_Movement_of_Liberia_for_Democracy_-
_Johnson_faction
[pem_pos=0, final_scores=0.40666634,
final_local_scores=0.55322397, pem=0.0, similarity=20.59902]
16: 21383 Nigeria
[pem_pos=0, final_scores=0.5174681,
final_local_scores=0.4908195, pem=-0.075027294,
similarity=19.144756]
17: 56823 Abuja
[pem_pos=0, final_scores=0.4020459,
final_local_scores=0.42675912, pem=0.0, similarity=16.825056]
18: 1618193
Economic_Community_of_West_African_States_Monitoring_Group
[pem_pos=0, final_scores=0.57418084,
final_local_scores=0.58958155, pem=0.0, similarity=21.684002]
19: 17791 Liberia
[pem_pos=0, final_scores=0.54915196,
final_local_scores=0.5020993, pem=-0.26536247,
similarity=20.51644]
20: 36848 Economic_Community_of_West_African_States
[pem_pos=0, final_scores=0.45298165,
final_local_scores=0.45013177, pem=0.0, similarity=17.522543]
```

Sample: b'77_Guinea'

Guinea calls two days of prayer . [1 1 CONAKRY 1] 1996-08-22
The West African state of Guinea declared Thursday and Friday
days of national prayer . A government statement , broadcast
repeatedly by state radio , said the two days of prayer were
" for the dead , for peace and prosperity in Guinea , the
victory of the new government and the health of the head of
state " . The precise reason for the call was not immediately
clear . Guinea 's president , [2 2 Lansana Conte 2] , vice-
president of the [3 3 Organisation of the Islamic Conference 3]
, left for [4 4 Kuwait 4] on August 16 to prepare the next [
5 5 OIC 5] summit in [6 6 Pakistan 6] in 1997 . [7 7 Koranic 7]
reading sessions and prayers were to be held in the farming
town of Badi-Tondon , near his home about 60 km (40 miles)
from the capital [8 8 Conakry 8] . Conte , an army general ,
survived a February army pay revolt which at the time he
described as a veiled attempt to topple him . He has since
named a prime minister for the first time since early in his
rule and ordered a crackdown on corruption . Conte seized
power in 1984 after the death of veteran [9 9 Marxist 9]
leader [10 10 Ahmed Sekou Toure 10] . He won elections in
1993 .

1: 56638 Conakry
2: 392412 Lansana_Conté
3: 47725 Organisation_of_Islamic_Cooperation
4: 7515890 Kuwait

```
5: 47725 Organisation_of_Islamic_Cooperation
6: 23235 Pakistan
7: 36922 Quran
8: 56638 Conakry
9: 1904053 Marxism
10: 391481 Ahmed_Sékou_Touré
1: 56638 Conakry
    [pem_pos=0, final_scores=0.38102776,
final_local_scores=0.39176124, pem=0.0, similarity=15.78065]
2: 392412 Lansana_Conté
    [pem_pos=0, final_scores=0.5176138,
final_local_scores=0.5809765, pem=0.0, similarity=21.42721]
3: 47725 Organisation_of_Islamic_Cooperation
    [pem_pos=0, final_scores=0.42655808,
final_local_scores=0.43965536, pem=0.0, similarity=17.209906]
4: 7515890 Kuwait
    [pem_pos=0, final_scores=0.40880007,
final_local_scores=0.40753376, pem=-0.3372397,
similarity=18.085295]
5: 47725 Organisation_of_Islamic_Cooperation
    [pem_pos=0, final_scores=0.5075055,
final_local_scores=0.52807814, pem=0.0, similarity=19.848618]
6: 23235 Pakistan
    [pem_pos=0, final_scores=0.54432315,
final_local_scores=0.575011, pem=-0.21944217,
similarity=22.442549]
7: 36922 Quran
    [pem_pos=0, final_scores=0.2651959,
final_local_scores=0.34628725, pem=0.0, similarity=14.423615]
8: 56638 Conakry
    [pem_pos=0, final_scores=0.29718834,
final_local_scores=0.30017933, pem=0.0, similarity=13.047663]
9: 1904053 Marxism
    [pem_pos=0, final_scores=0.32374698,
final_local_scores=0.390276, pem=0.0, similarity=15.736328]
10: 391481 Ahmed_Sékou_Touré
    [pem_pos=0, final_scores=0.49042708,
final_local_scores=0.52467775, pem=0.0, similarity=19.747143]
```

Sample: b'78_South'

```
[1 1 South African 1] answers [2 2 U.S. 2] message in a
bottle . [3 3 JOHANNESBURG 3] 1996-08-22 A [4 4 South African
4] boy is writing back to an [5 5 American 5] girl whose
message in a bottle he found washed up on President [6 6 Nels
on Mandela 6] 's old prison island . But Carlo Hoffmann , an
11-year-old jailer 's son who found the bottle on the beach
at Robben Island off [7 7 Cape Town 7] after winter storms ,
will send his letter back by ordinary mail on Thursday , the
post office said . It will be sent for free . Danielle Murray
from [8 8 Sandusky 8] , [9 9 Ohio 9] , the same age as her
new penfriend , asked for a reply from whoever received the
message she flung on its journey months ago on the other side
```

```
of the Atlantic Ocean .
1: 17416221 South_Africa
2: 3434750 United_States
3: 58893 Johannesburg
4: 17416221 South_Africa
5: 3434750 United_States
6: 21492751 Nelson_Mandela
7: 6653 Cape_Town
8: 14301875 Sandusky,_Ohio
9: 22199 Ohio
1: 17416221 South_Africa
    [pem_pos=0, final_scores=0.48750174,
final_local_scores=0.46626866, pem=-0.1389451,
similarity=18.759705]
2: 3434750 United_States
    [pem_pos=0, final_scores=0.5034134,
final_local_scores=0.47662252, pem=-0.14075686,
similarity=19.078539]
3: 58893 Johannesburg
    [pem_pos=0, final_scores=0.49193555,
final_local_scores=0.47027338, pem=-0.09562059,
similarity=18.643608]
4: 17416221 South_Africa
    [pem_pos=0, final_scores=0.5409519,
final_local_scores=0.5246548, pem=-0.1389451,
similarity=20.502064]
5: 3434750 United_States
    [pem_pos=0, final_scores=0.45886177,
final_local_scores=0.42795664, pem=-0.12994486,
similarity=17.567453]
6: 21492751 Nelson_Mandela
    [pem_pos=0, final_scores=0.40830815,
final_local_scores=0.45296246, pem=0.0, similarity=17.607016]
7: 6653 Cape_Town
    [pem_pos=0, final_scores=0.44976825,
final_local_scores=0.41845548, pem=0.0, similarity=16.57726]
8: 14301875 Sandusky,_Ohio
    [pem_pos=0, final_scores=0.3080594,
final_local_scores=0.3433983, pem=-0.61382115,
similarity=17.675457]
9: 22199 Ohio
    [pem_pos=0, final_scores=0.67160696,
final_local_scores=0.6920876, pem=-0.19147073,
similarity=25.784235]
```

Sample: b'79_Rottweiler'
[1 1 Rottweiler 1] kills [2 2 South African 2] toddler . [3 3
JOHANNESBURG 3] 1996-08-22 A rottweiler dog belonging to an
elderly [4 4 South African 4] couple savaged to death their
two-year-old grandson who was visiting , police said on
Thursday . The dog attacked Louis Booy in the front garden of
his grandparents ' house in [5 5 Vanderbijlpark 5] near [6 6 J

ohannesburg 6] on Tuesday . His bloody body was lying in the garden when his parents arrived in the afternoon to pick him up . It was unclear where the grandparents were at the time . Dogs fierce enough to scare off burglars are becoming increasingly popular in the crime-infested [7 7 Johannesburg 7] area .

```
1: 271052 Rottweiler
2: 17416221 South_Africa
3: 58893 Johannesburg
4: 17416221 South_Africa
5: 1413961 Vanderbijlpark
6: 58893 Johannesburg
7: 58893 Johannesburg
1: 271052 Rottweiler
    [pem_pos=0, final_scores=0.40176356,
final_local_scores=0.5699509, pem=0.0, similarity=21.098185]
2: 17416221 South_Africa
    [pem_pos=0, final_scores=0.48054677,
final_local_scores=0.4467107, pem=-0.1389451,
similarity=18.176056]
3: 58893 Johannesburg
    [pem_pos=0, final_scores=0.5568386,
final_local_scores=0.50872165, pem=-0.09562059,
similarity=19.790981]
4: 17416221 South_Africa
    [pem_pos=0, final_scores=0.4509341,
final_local_scores=0.41436338, pem=-0.1389451,
similarity=17.210747]
5: 1413961 Vanderbijlpark
    [pem_pos=0, final_scores=0.47823185,
final_local_scores=0.52959836, pem=0.0, similarity=19.893984]
6: 58893 Johannesburg
    [pem_pos=0, final_scores=0.5263242,
final_local_scores=0.47538942, pem=-0.09562059,
similarity=18.796282]
7: 58893 Johannesburg
    [pem_pos=0, final_scores=0.533762,
final_local_scores=0.483514, pem=-0.09562059,
similarity=19.038736]
```

Sample: b'80_INDICATORS'

```
INDICATORS - [1 1 Hungary 1] - updated Aug 22 . [2 2 BUDAPEST
2] 1996-08-22 The latest indicators : CPI ( pct ) July
+0.4m / m ; 23.0yr / yr ( June +0.9;+23.6 ) PPI ( pct ) June
+0.7 m / m;+21.5yr / yr ( May +1.7;+22.0 ) Industry output
( pct ) June - 7.8 m / m;-0.2yr / yr ( May +7.3;-3.6 )
Current account Jan-May - $ 738 million ( Jan-April - $ 748
million ) NBH trade balance Jan-May - $ 934 million ( Jan-
April - $ 774 million ) MIT trade balance Jan-June - $ 1.45
bln ( Jan-May - $ 1.24 bln ) Gross foreign debt May $ 27,246.5
million ( April $ 28,716.8 million ) Net foreign
debt May $ 14,390.7 million ( April $ 15,704.3 million )
```

Unemployment (pct) July 10.8 pct (June 10.6 pct) Budget deficit (HUF) Jan-July 102 bln (Jan-June 122 bln) T-bill yields % (1mo) 22.95 (3mo) 23.02 (6mo) 23.53 (1yr) 24.40 Government bond yields : (2-yr 1998 / J) 25.49, (3-yr 1999 / c) 24.44 The NBH is BBB-minus by [3 3 Duff & Phelps 3] , IBCA and Thomson BankWatch , BB-plus by [4 4 S&P 4] , BA1 by Moody 's Investors Service , BBB+ by the Japan Credit Rating Agency . The NBH trade data is based on cash flow , MIT data on customs statistics . -- [5 5 Budapest 5] newsroom (36 1)266 2410

1: 13275 Hungary

2: 36787 Budapest

3: 24565776 Duff_&_Phelps

4: 164386 Standard_&_Poor's

5: 36787 Budapest

1: 13275 Hungary
[pem_pos=0, final_scores=0.5960483,
final_local_scores=0.611243, pem=-0.32577527,
similarity=24.102041]

2: 36787 Budapest
[pem_pos=0, final_scores=0.5427385,
final_local_scores=0.5345177, pem=-0.030229695,
similarity=20.205181]

3: 24565776 Duff_&_Phelps
[pem_pos=0, final_scores=0.46568942,
final_local_scores=0.6099982, pem=0.0, similarity=22.293276]

4: 164386 Standard_&_Poor's
[pem_pos=0, final_scores=0.47750288,
final_local_scores=0.59120506, pem=0.0, similarity=21.73245]

5: 36787 Budapest
[pem_pos=0, final_scores=0.5677946,
final_local_scores=0.56188774, pem=-0.030229695,
similarity=21.021957]

Sample: b'81_Fifty'

Fifty [1 1 Russians 1] die in clash with rebels-Interfax . [2 2 MOSCOW 2] 1996-08-22 At least 50 [3 3 Russian 3] servicemen have been killed in a battle with separatist rebels which erupted in the [1 Chechen 1] capital [4 4 Grozny 4] on Thursday and continued after [5 5 Russia 5] and the rebels agreed a truce , [6 6 Interfax 6] news agency said . [7 7 Interfax 7] quoted [8 8 Russian 8] military command in [9 9 Chechnya 9] as saying that about 200 [10 10 interior ministry 10] forces , sent on reconnaissance mission , clashed with rebels at Minutka Square . The [11 11 Interfax 11] report could not be independently confirmed . [12 12 Moscow 12] peacemaker [13 13 Alexander Lebed 13] and rebel chief-of-staff [14 14 Aslan Maskhadov 14] signed an agreement earlier on Thursday under which the two sides would cease all hostilities at noon (0800 [15 15 GMT 15]) on Friday . [16 16 Interfax 16] made clear that the [17 17 interior ministry 17] detachment had been sent on the mission before the truce deal had been

signed at the local equivalent of 1500 [18 18 GMT 18] . But fierce fighting still raged at 1600 [19 19 GMT 19] , [20 20 Interfax 20] said . It quoted a source in the [21 21 Russian 21] command in [22 22 Chechnya 22] as saying that the servicemen were outnumbered by the rebels .

1: 25391 Russia
2: 19004 Moscow
3: 25391 Russia
4: 140790 Grozny
5: 25391 Russia
6: 4921404 Interfax
7: 4921404 Interfax
8: 25391 Russia
9: 6095 Chechnya
10: 354620 Interior_ministry
11: 4921404 Interfax
12: 19004 Moscow
13: 431579 Alexander_Lebed
14: 340963 Aslan_Maskhadov
15: 12701 Greenwich_Mean_Time
16: 4921404 Interfax
17: 354620 Interior_ministry
18: 12701 Greenwich_Mean_Time
19: 12701 Greenwich_Mean_Time
20: 4921404 Interfax
21: 25391 Russia
22: 6095 Chechnya
1: 25391 Russia
[pem_pos=1, final_scores=0.7580269,
final_local_scores=0.76672274, pem=-1.8802736,
similarity=37.19547]
2: 19004 Moscow
[pem_pos=0, final_scores=0.68708915,
final_local_scores=0.6898879, pem=-0.106582455,
similarity=25.256956]
3: 25391 Russia
[pem_pos=1, final_scores=0.87903804,
final_local_scores=0.89890903, pem=-1.0515022,
similarity=36.633186]
4: 140790 Grozny
[pem_pos=0, final_scores=0.47464818,
final_local_scores=0.4712308, pem=0.0, similarity=18.15218]
5: 25391 Russia
[pem_pos=0, final_scores=0.8322049,
final_local_scores=0.84775084, pem=-0.20299491,
similarity=30.492207]
6: 4921404 Interfax
[pem_pos=0, final_scores=0.71287066,
final_local_scores=0.7328829, pem=0.0, similarity=25.960402]
7: 4921404 Interfax
[pem_pos=0, final_scores=0.5394328,
final_local_scores=0.5434284, pem=0.0, similarity=20.306702]

8: 25391 Russia
[pem_pos=1, final_scores=0.88955706,
final_local_scores=0.9103994, pem=-1.0515022,
similarity=36.97608]
9: 6095 Chechnya
[pem_pos=0, final_scores=0.5520945,
final_local_scores=0.5342161, pem=0.0, similarity=20.031788]
10: 354620 Interior_ministry
[pem_pos=0, final_scores=0.55466515,
final_local_scores=0.592095, pem=-0.57122487,
similarity=24.865417]
11: 4921404 Interfax
[pem_pos=0, final_scores=0.4760918,
final_local_scores=0.47423798, pem=0.0, similarity=18.24192]
12: 19004 Moscow
[pem_pos=0, final_scores=0.41024303,
final_local_scores=0.38747555, pem=-0.106582455,
similarity=16.232368]
13: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.7444168,
final_local_scores=0.8039676, pem=0.0, similarity=28.081709]
14: 340963 Aslan_Maskhadov
[pem_pos=0, final_scores=0.55742246,
final_local_scores=0.5946844, pem=0.0, similarity=21.836283]
15: 12701 Greenwich_Mean_Time
[pem_pos=0, final_scores=0.6892094,
final_local_scores=0.774003, pem=0.0, similarity=27.187508]
16: 4921404 Interfax
[pem_pos=0, final_scores=0.52849734,
final_local_scores=0.53148305, pem=0.0, similarity=19.950228]
17: 354620 Interior_ministry
[pem_pos=0, final_scores=0.43981487,
final_local_scores=0.46663862, pem=-0.57122487,
similarity=21.12155]
18: 12701 Greenwich_Mean_Time
[pem_pos=0, final_scores=0.6273405,
final_local_scores=0.7064206, pem=0.0, similarity=25.170713]
19: 12701 Greenwich_Mean_Time
[pem_pos=0, final_scores=0.6327216,
final_local_scores=0.7122987, pem=0.0, similarity=25.346127]
20: 4921404 Interfax
[pem_pos=0, final_scores=0.53724766,
final_local_scores=0.54104143, pem=0.0, similarity=20.235468]
21: 25391 Russia
[pem_pos=1, final_scores=0.8561949,
final_local_scores=0.87395626, pem=-1.0515022,
similarity=35.888546]
22: 6095 Chechnya
[pem_pos=0, final_scores=0.5244804,
final_local_scores=0.5040519, pem=0.0, similarity=19.131628]
1: 6095 Chechnya
[pem_pos=2, final_scores=0.41649073,

```
final_local_scores=0.38608962, pem=-0.8617233,  
similarity=20.297583]
```

Sample: b'82_Polish'
[1 1 Polish 1] schoolgirl blackmailer wanted textbooks . [2 2
GDANSK 2] , [3 3 Poland 3] 1996-08-22 A [4 4 Polish 4]
schoolgirl blackmailed two women with anonymous letters
threatening death and later explained that she needed money
for textbooks , police said on Thursday . " The 13-year-old
girl tried to extract 60 and 70 zlotys (\$ 22 and \$ 26) from
two residents of [5 5 Sierakowice 5] by threatening to take
their lives , " a police spokesman said in the nearby
northern city of [6 6 Gdańsk 6] on Thursday . He said the
women reported the blackmail letters and police caught the
girl on Wednesday as she tried to pick up the cash at the [7 7
Sierakowice 7] railway station . " Interviewed in the
presence of a psychologist , she said she wanted to use the
money for school books and clothes , " spokesman Kazimierz
Socha told [8 8 Reuters 8] . He said the case of the girl ,
from a poor family that had never been in trouble with the
law , would go before a special court dealing with underage
offenders .

```
1: 22936 Poland  
2: 12099 Gdańsk  
3: 22936 Poland  
4: 22936 Poland  
5: 21666402 Sierakowice,_Pomeranian_Voivodeship  
6: 12099 Gdańsk  
7: 21666402 Sierakowice,_Pomeranian_Voivodeship  
8: 18998750 Reuters  
1: 22936 Poland  
    [pem_pos=0, final_scores=0.6883318,  
final_local_scores=0.67527395, pem=-0.81518245,  
similarity=28.674322]  
2: 12099 Gdańsk  
    [pem_pos=0, final_scores=0.48752862,  
final_local_scores=0.46017128, pem=0.0, similarity=17.822142]  
3: 22936 Poland  
    [pem_pos=0, final_scores=0.5641574,  
final_local_scores=0.5396324, pem=-0.15333635,  
similarity=21.027287]  
4: 22936 Poland  
    [pem_pos=0, final_scores=0.66066974,  
final_local_scores=0.6450574, pem=-0.81518245,  
similarity=27.7726]  
5: 21666402 Sierakowice,_Pomeranian_Voivodeship  
    [pem_pos=0, final_scores=0.4729455,  
final_local_scores=0.5532118, pem=-0.2972515,  
similarity=22.215157]  
6: 12099 Gdańsk  
    [pem_pos=0, final_scores=0.4221576,  
final_local_scores=0.38876337, pem=0.0, similarity=15.691187]
```

```
7: 21666402 Sierakowice,_Pomeranian_Voivodeship
    [pem_pos=0, final_scores=0.4223094,
final_local_scores=0.4978996, pem=-0.2972515,
similarity=20.56453]
8: 18998750 Reuters
    [pem_pos=0, final_scores=0.30898362,
final_local_scores=0.4455455, pem=0.0, similarity=17.38568]
```

Sample: b'83_Czech'

```
[1 1 Czech 1] CNB-120 index rises 1.2 pts to 869.3 . [2 2 PR
AGUE 2] 1996-08-22 The CNB-120 index , a broad daily measure
of [3 3 Czech 3] equities , rose 1.2 points on Thursday to
869.3 , the [4 4 Czech National Bank 4] ( CNB ) said . Eight
of the ten sectoral indices rose , with the banking index
rising the most , up 14.4 points to 1,294.5 . -- Prague
Newsroom , 42-2-2423-0003
```

```
1: 5321 Czech_Republic
```

```
2: 23844 Prague
```

```
3: 5321 Czech_Republic
```

```
4: 1683932 Czech_National_Bank
```

```
1: 5321 Czech_Republic
```

```
    [pem_pos=1, final_scores=0.49445635,
final_local_scores=0.4559672, pem=-1.208016,
similarity=24.266062]
```

```
2: 23844 Prague
```

```
    [pem_pos=0, final_scores=0.4515291,
final_local_scores=0.450912, pem=-0.05392049,
similarity=17.839054]
```

```
3: 5321 Czech_Republic
```

```
    [pem_pos=1, final_scores=0.5458292,
final_local_scores=0.5120842, pem=-1.208016,
similarity=25.940704]
```

```
4: 1683932 Czech_National_Bank
```

```
    [pem_pos=0, final_scores=0.31647414,
final_local_scores=0.44962108, pem=0.0, similarity=17.507303]
```

Sample: b'84_Russians'

```
[1 1 Russians 1] , rebels sign deal in [2 2 Chechnya 2] . [3
3 NOVYE ATAGI 3] , [4 4 Russia 4] 1996-08-22 [5 5 Russian 5]
President [6 6 Boris Yeltsin 6] 's security supremo [7 7 Alex
ander Lebed 7] and [8 8 Chechen 8] rebel chief-of-staff [9 9 A
slan Maskhadov 9] signed a deal on Thursday aimed at ending
three weeks of renewed fighting in the region . The final
contents of the document negotiated in this village south of
the [10 10 Chechen 10] capital [11 11 Grozny 11] have not
been officially disclosed . [12 12 Itar-Tass 12] news agency
said it provided for the disengagement of [13 13 Russian 13]
and rebel forces in [14 14 Chechnya 14] .
```

```
1: 25391 Russia
```

```
2: 6095 Chechnya
```

```
3: 4813770 Novye_Atagi
```

```
4: 25391 Russia
```

5: 25391 Russia
6: 23278520 Boris_Yeltsin
7: 431579 Alexander_Lebed
8: 6095 Chechnya
9: 340963 Aslan_Maskhadov
10: 6095 Chechnya
11: 140790 Grozny
12: 434852 Information_Telegraph_Agency_of_Russia
13: 25391 Russia
14: 6095 Chechnya
1: 25391 Russia
[pem_pos=1, final_scores=0.65107274,
final_local_scores=0.64888924, pem=-1.8802736,
similarity=33.67908]
2: 6095 Chechnya
[pem_pos=0, final_scores=0.5910986,
final_local_scores=0.54956174, pem=0.0, similarity=20.48973]
3: 4813770 Novye_Atagi
[pem_pos=0, final_scores=0.43352807,
final_local_scores=0.519001, pem=0.0, similarity=19.577738]
4: 25391 Russia
[pem_pos=0, final_scores=0.8456853,
final_local_scores=0.861474, pem=-0.20299491,
similarity=30.901733]
5: 25391 Russia
[pem_pos=1, final_scores=0.84701425,
final_local_scores=0.86292565, pem=-1.0515022,
similarity=35.55937]
6: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.61661726,
final_local_scores=0.63567674, pem=0.0, similarity=23.059574]
7: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.7042034,
final_local_scores=0.7568922, pem=0.0, similarity=26.676886]
8: 6095 Chechnya
[pem_pos=2, final_scores=0.5704087,
final_local_scores=0.52696115, pem=-0.8617233,
similarity=24.50147]
9: 340963 Aslan_Maskhadov
[pem_pos=0, final_scores=0.58725625,
final_local_scores=0.6074008, pem=0.0, similarity=22.215763]
10: 6095 Chechnya
[pem_pos=2, final_scores=0.5128512,
final_local_scores=0.46408826, pem=-0.8617233,
similarity=22.625217]
11: 140790 Grozny
[pem_pos=0, final_scores=0.48173982,
final_local_scores=0.46684152, pem=0.0, similarity=18.021194]
12: 434852 Information_Telegraph_Agency_of_Russia
[pem_pos=0, final_scores=0.65197617,
final_local_scores=0.7210241, pem=0.0, similarity=25.60651]
13: 25391 Russia

```
[pem_pos=2, final_scores=0.9369741,
final_local_scores=0.961193, pem=-1.0515022,
similarity=38.491863]
14: 6095 Chechnya
    [pem_pos=0, final_scores=0.59043455,
final_local_scores=0.54883635, pem=0.0, similarity=20.468084]

Sample: b'85 Lebed'
[1 1 Lebed 1] aide says Russian-Chechen talks going well . [2
2 NOVYE ATAGI 2] , [3 3 Russia 3] 1996-08-22 Talks between [4
4 Russia 4] 's [5 5 Alexander Lebed 5] and [6 6 Chechen 6]
separatist leaders were going well on Thursday and the two
sides were working out a detailed schedule on how to stop the
war , a [7 7 Lebed 7] aide said . Press spokesman Alexander
Barkhatov told reporters the negotiations , being held at
this rebel-held village some 20 km ( 12 miles ) south of the [8
8 Chechen 8] capital [9 9 Grozny 9] , were progressing
briskly and being conducted in a good mood . He said a
document would be completed in an hour 's time for signature
by the two sides , who were working on a " day-by-day
schedule to stop the war in [10 10 Chechnya 10] . "
1: 431579 Alexander_Lebed
2: 4813770 Novye_Atagi
3: 25391 Russia
4: 25391 Russia
5: 431579 Alexander_Lebed
6: 6095 Chechnya
7: 431579 Alexander_Lebed
8: 6095 Chechnya
9: 140790 Grozny
10: 6095 Chechnya
1: 431579 Alexander_Lebed
    [pem_pos=0, final_scores=0.60457623,
final_local_scores=0.59628886, pem=0.0, similarity=21.884161]
2: 4813770 Novye_Atagi
    [pem_pos=0, final_scores=0.40110755,
final_local_scores=0.4910062, pem=0.0, similarity=18.742317]
3: 25391 Russia
    [pem_pos=0, final_scores=0.7888784,
final_local_scores=0.84181136, pem=-0.20299491,
similarity=30.314959]
4: 25391 Russia
    [pem_pos=0, final_scores=0.8377579,
final_local_scores=0.89520484, pem=-0.20299491,
similarity=31.908329]
5: 431579 Alexander_Lebed
    [pem_pos=0, final_scores=0.6953742,
final_local_scores=0.6833327, pem=0.0, similarity=24.481724]
6: 6095 Chechnya
    [pem_pos=2, final_scores=0.53535825,
final_local_scores=0.51014394, pem=-0.8617233,
similarity=23.99961]
```

```
7: 431579 Alexander_Lebed
    [pem_pos=0, final_scores=0.6142698,
final_local_scores=0.6068776, pem=0.0, similarity=22.200151]
8: 6095 Chechnya
    [pem_pos=2, final_scores=0.47661656,
final_local_scores=0.44597757, pem=-0.8617233,
similarity=22.084759]
9: 140790 Grozny
    [pem_pos=0, final_scores=0.47101432,
final_local_scores=0.46452802, pem=0.0, similarity=17.952156]
10: 6095 Chechnya
    [pem_pos=0, final_scores=0.5504542,
final_local_scores=0.526634, pem=0.0, similarity=19.805521]
```

Sample: b'86_Yeltsin'

[1 1 Yeltsin 1] shown on [2 2 Russian 2] television . [3 3 MO SCOW 3] 1996-08-22 [4 4 Russian 4] television showed a brief clip of [5 5 Boris Yeltsin 5] on Thursday , with the president laughing and smiling as he spoke to nominee health minister Tatyana Dmitrieva . It was the first time the president had been shown on television since he was inaugurated for a second term in office on August 9 . He returned to the [6 6 Kremlin 6] on Thursday after a two-day break in the lakelands of northwestern [7 7 Russia 7] .

```
1: 23278520 Boris_Yeltsin
2: 25391 Russia
3: 19004 Moscow
4: 25391 Russia
5: 23278520 Boris_Yeltsin
6: 381075 Moscow_Kremlin
7: 25391 Russia
1: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.45785666,
final_local_scores=0.4524877, pem=0.0, similarity=17.592848]
2: 25391 Russia
    [pem_pos=1, final_scores=0.79537374,
final_local_scores=0.7882079, pem=-1.0515022,
similarity=33.32964]
3: 19004 Moscow
    [pem_pos=0, final_scores=0.6895359,
final_local_scores=0.68164486, pem=-0.106582455,
similarity=25.010967]
4: 25391 Russia
    [pem_pos=1, final_scores=0.85692495,
final_local_scores=0.85544324, pem=-1.0515022,
similarity=35.33608]
5: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.55278367,
final_local_scores=0.5561811, pem=0.0, similarity=20.687265]
6: 381075 Moscow_Kremlin
    [pem_pos=1, final_scores=0.38184196,
final_local_scores=0.40027344, pem=-0.76193243,
```

```
similarity=20.178177]
7: 25391 Russia
    [pem_pos=0, final_scores=0.8023488,
final_local_scores=0.79582703, pem=-0.20299491,
similarity=28.942696]

Sample: b'87_PRESS'
PRESS DIGEST - Bosnia - Aug 22 . [1 1 SARAJEVO 1] 1996-08-22
These are the leading stories in the [2 2 Sarajevo 2] press
on Thursday . [3 3 Reuters 3] has not verified these stories
and does not vouch for their accuracy . [4 4 OSLOBODJENJE 4]
- The Bosnian federation launches a common payment system on
Friday . Under the new system taxes and customs may be paid
in the Bosnian dinar , the [5 5 Croatian 5] kuna or the
Deutsche mark until a new Bosnian currency is introduced . -
The president of the Bosnian Association for Refugees and
Displaced Persons , Mirhunisa Komarica says many survivors of
the 1995 massacre in the Bosnian town of [6 6 Srebrenica 6]
are languishing as forced laborers in [7 7 Serbian 7] mines .
According to Komarica , 2,400 male residents of [8 8 Srebreni
ca 8] work in the Trepca mine and 1,900 work in a mine in
Aleksandrovac . [9 9 DNEVNI AVAZ 9] - [10 10 Slovenian 10]
police briefly detain two Bosnian opposition leaders in [11 11
Ljubljana 11] and cancel opposition political rallies in [12 1
2 Ljubljana 12] and [13 13 Maribor 13] . -- [14 14 Sarajevo 14
] newsroom , +387-71-663-864 .
1: 26786 Sarajevo
2: 26786 Sarajevo
3: 18998750 Reuters
4: 751324 Oslobođenje
5: 5573 Croatia
6: 28912 Srebrenica
7: 29265 Serbia
8: 28912 Srebrenica
9: 751326 Dnevni_avaz
10: 27338 Slovenia
11: 42374 Ljubljana
12: 42374 Ljubljana
13: 20670 Maribor
14: 26786 Sarajevo
1: 26786 Sarajevo
    [pem_pos=0, final_scores=0.44748992,
final_local_scores=0.39020276, pem=-0.0017846047,
similarity=15.743847]
2: 26786 Sarajevo
    [pem_pos=0, final_scores=0.43586105,
final_local_scores=0.37750006, pem=-0.0017846047,
similarity=15.364772]
3: 18998750 Reuters
    [pem_pos=0, final_scores=0.26174414,
final_local_scores=0.3378242, pem=0.0, similarity=14.171061]
4: 751324 Oslobođenje
```

```
[pem_pos=0, final_scores=0.5209202,
final_local_scores=0.6612044, pem=0.0, similarity=23.821371]
5: 5573 Croatia
    [pem_pos=1, final_scores=0.47068077,
final_local_scores=0.4376828, pem=-1.1038951,
similarity=23.154194]
6: 28912 Srebrenica
    [pem_pos=0, final_scores=0.2820981,
final_local_scores=0.30280304, pem=0.0, similarity=13.125959]
7: 29265 Serbia
    [pem_pos=1, final_scores=0.58963746,
final_local_scores=0.56924814, pem=-1.2504034,
similarity=27.877102]
8: 28912 Srebrenica
    [pem_pos=0, final_scores=0.3208242,
final_local_scores=0.34510535, pem=0.0, similarity=14.388345]
9: 751326 Dnevni_avaz
    [pem_pos=0, final_scores=0.46170378,
final_local_scores=0.57278097, pem=0.0, similarity=21.18264]
10: 27338 Slovenia
    [pem_pos=0, final_scores=0.3566472,
final_local_scores=0.32353753, pem=-0.5440836,
similarity=16.70353]
11: 42374 Ljubljana
    [pem_pos=0, final_scores=0.463086,
final_local_scores=0.4207986, pem=0.0, similarity=16.647182]
12: 42374 Ljubljana
    [pem_pos=0, final_scores=0.49885678,
final_local_scores=0.45987272, pem=0.0, similarity=17.813232]
13: 20670 Maribor
    [pem_pos=0, final_scores=0.40751487,
final_local_scores=0.37664527, pem=-0.01734017,
similarity=15.423859]
14: 26786 Sarajevo
    [pem_pos=0, final_scores=0.47292024,
final_local_scores=0.41798156, pem=-0.0017846047,
similarity=16.57282]
```

Sample: b'88_Grozny'

[1 1 Grozny 1] quiet overnight after raids . ALKHAN-YURT , [2 2 Russia 2] 1996-08-22 The city of [3 3 Grozny 3] , pounded by [4 4 Russian 4] planes and artillery for hours on Wednesday , calmed down overnight , although sporadic explosions and shooting could still be heard . [5 5 Reuters 5] correspondent Lawrence Sheets , speaking from the nearby village of Alkhan-Yurt , said he had heard little from [6 6 Gr ozny 6] since Wednesday evening 's arrival of [7 7 Russian 7] security chief [8 8 Alexander Lebed 8] , who said he " came with peace " . A couple of helicopters flew over the city early on Thursday morning , but did not appear to be firing at anything . [9 9 Lebed 9] said on Wednesday he had clinched a truce with [10 10 Chechen 10] separatists and he promised

to halt a threatened bombing assault on [11 11 Grozny 11] , which the rebels have held since August 6 .

1: 140790 Grozny
2: 25391 Russia
3: 140790 Grozny
4: 25391 Russia
5: 18998750 Reuters
6: 140790 Grozny
7: 25391 Russia
8: 431579 Alexander_Lebed
9: 431579 Alexander_Lebed
10: 6095 Chechnya
11: 140790 Grozny
1: 140790 Grozny
[pem_pos=0, final_scores=0.4570312, final_local_scores=0.41381383, pem=0.0, similarity=16.438744]
2: 25391 Russia
[pem_pos=0, final_scores=0.8264912, final_local_scores=0.8566713, pem=-0.20299491, similarity=30.75841]
3: 140790 Grozny
[pem_pos=0, final_scores=0.67860603, final_local_scores=0.6558507, pem=0.0, similarity=23.661606]
4: 25391 Russia
[pem_pos=1, final_scores=0.85705155, final_local_scores=0.8900539, pem=-1.0515022, similarity=36.36893]
5: 18998750 Reuters
[pem_pos=0, final_scores=0.33106542, final_local_scores=0.39366966, pem=0.0, similarity=15.837602]
6: 140790 Grozny
[pem_pos=0, final_scores=0.60284084, final_local_scores=0.57308877, pem=0.0, similarity=21.191824]
7: 25391 Russia
[pem_pos=1, final_scores=0.90824014, final_local_scores=0.94596964, pem=-1.0515022, similarity=38.037567]
8: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.7969863, final_local_scores=0.8301601, pem=0.0, similarity=28.863346]
9: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.64935035, final_local_scores=0.6879287, pem=0.0, similarity=24.618877]
10: 6095 Chechnya
[pem_pos=2, final_scores=0.5277116, final_local_scores=0.5077822, pem=-0.8617233, similarity=23.929132]
11: 140790 Grozny
[pem_pos=0, final_scores=0.5791016, final_local_scores=0.5471572, pem=0.0, similarity=20.417976]

Sample: b'89_Boat'

Boat passengers rescued off [1 1 Colombian 1] coast . [2 2 B0 GOTA 2] , [3 3 Colombia 3] 1996-08-22 [4 4 Colombia 4] 's [5 5 Coast Guard 5] on Thursday rescued 12 people lost for three days in an open boat off the Pacific coast , officials said . The boat had been missing since Monday afternoon when it left the tiny island of [6 6 Gorgona 6] off [7 7 Colombia 7] 's southwest coast with sightseers for a return trip to Narino province , near the border with [8 8 Ecuador 8] . The boat ran out of fuel and did not have a radio to call for help , Navy spokesman Lt. Italo Pineda said . He said 11 passengers and one boatman survived on coconuts and rainwater during 65 hours lost at sea . The boat was towed to the port city of [9 9 Buenaventura 9] .

1: 5222 Colombia
2: 211271 Bogotá
3: 5222 Colombia
4: 5222 Colombia
5: 146790 Coast_guard
6: 5418443 Gorgona_Island,_Colombia
7: 5222 Colombia
8: 9334 Ecuador
9: 7996627 Buenaventura,_Valle_del_Cauca
1: 5222 Colombia
[pem_pos=0, final_scores=0.49460477,
final_local_scores=0.4459538, pem=-0.32036036,
similarity=19.140032]
2: 211271 Bogotá
[pem_pos=0, final_scores=0.5902614,
final_local_scores=0.60881853, pem=0.0, similarity=22.258072]
3: 5222 Colombia
[pem_pos=0, final_scores=0.61260796,
final_local_scores=0.5748543, pem=-0.22328302,
similarity=22.45876]
4: 5222 Colombia
[pem_pos=0, final_scores=0.73402584,
final_local_scores=0.7074849, pem=-0.22328302,
similarity=26.416721]
5: 146790 Coast_guard
[pem_pos=1, final_scores=0.1607114,
final_local_scores=0.24967411, pem=-1.3230537,
similarity=18.735458]
6: 5418443 Gorgona_Island,_Colombia
[pem_pos=1, final_scores=0.30013222,
final_local_scores=0.35480085, pem=-0.33647218,
similarity=16.507465]
7: 5222 Colombia
[pem_pos=0, final_scores=0.6749941,
final_local_scores=0.6430017, pem=-0.22328302,
similarity=24.492413]
8: 9334 Ecuador
[pem_pos=0, final_scores=0.69002205,
final_local_scores=0.6923945, pem=-0.21404564,

```
similarity=25.916159]
9: 7996627 Buenaventura,_Valle_del_Cauca
    [pem_pos=0, final_scores=0.36172396,
final_local_scores=0.39342117, pem=0.0, similarity=15.830185]
```

Sample: b'90_Argentine'

```
[1 1 Argentine 1] July raw steel output up 14.8 pct vs '95 . [2 2 BUENOS AIRES 2] 1996-08-22 [3 3 Argentine 3] raw steel output was 355,900 tonnes in July , 14.8 percent higher than in July 1995 and up 1.9 percent from June , Steel Industry Center said Thursday . Primary iron output was 297,700 tonnes , 14.5 percent more than last July and 0.1 percent more than in June . Hot laminate production was 349,000 tonnes , 3.2 percent up from July 1995 and 0.8 percent up from June . Production of cold laminates was 120,500 tonnes , 4.2 percent higher than the same month last year and 11 percent higher than in June . -- Jason Webb , Buenos Aires Newsroom +541 318-0655
```

```
1: 18951905 Argentina
2: 255919 Buenos_Aires
3: 18951905 Argentina
1: 18951905 Argentina
    [pem_pos=0, final_scores=0.65431386,
final_local_scores=0.6017277, pem=-0.18885723,
similarity=23.073503]
2: 255919 Buenos_Aires
    [pem_pos=0, final_scores=0.39930272,
final_local_scores=0.3672738, pem=-0.12597719,
similarity=15.734979]
3: 18951905 Argentina
    [pem_pos=0, final_scores=0.7618664,
final_local_scores=0.71921253, pem=-0.18885723,
similarity=26.579485]
```

Sample: b'91_Peru'

```
[1 1 Peru 1] 's guerrillas kill one , take 8 hostage in jungle . [2 2 LIMA 2] , [3 3 Peru 3] 1996-08-21 [4 4 Peruvian 4] guerrillas killed one man and took eight people hostage after taking over a village in the country 's northeastern jungle region , anti- terrorist police sources said on Wednesday . For three hours on Tuesday , around 100 members of the Maoist rebel group [5 5 Shining Path 5] took control of Alomella Robles , a small village about 345 miles ( 550 km ) northeast of [6 6 Lima 6] , the sources said . Some guerrillas made villagers listen to propaganda speeches in the village centre , others forced passing motorists out of their cars and daubed their vehicles with slogans . By Wednesday , the whereabouts of the eight hostages was still not known , the sources said . In recent months the [7 7 Shining Path 7] , severely weakened since the 1992 capture of its leader [8 8 Abimael Guzman 8] , has been stepping up both its military and propaganda activities . [9 9 Peru 9] 's
```

guerrilla conflicts have cost at least 30,000 lives and \$ 25 billion in damage to infrastructure since 1980 .

1: 170691 Peru
2: 85423 Lima
3: 170691 Peru
4: 170691 Peru
5: 28521 Shining_Path
6: 85423 Lima
7: 28521 Shining_Path
8: 481609 Abimael_Guzmán
9: 170691 Peru
1: 170691 Peru
[pem_pos=0, final_scores=0.6492438,
final_local_scores=0.63438606, pem=-0.23383401,
similarity=24.292685]
2: 85423 Lima
[pem_pos=0, final_scores=0.5217635,
final_local_scores=0.5102635, pem=-0.32765645,
similarity=21.09884]
3: 170691 Peru
[pem_pos=0, final_scores=0.7140059,
final_local_scores=0.7051288, pem=-0.23383401,
similarity=26.403788]
4: 170691 Peru
[pem_pos=0, final_scores=0.7525367,
final_local_scores=0.7472179, pem=-0.22287759,
similarity=27.600227]
5: 28521 Shining_Path
[pem_pos=0, final_scores=0.61673975,
final_local_scores=0.6463636, pem=0.0, similarity=23.378492]
6: 85423 Lima
[pem_pos=0, final_scores=0.44838744,
final_local_scores=0.43011123, pem=-0.32765645,
similarity=18.706936]
7: 28521 Shining_Path
[pem_pos=0, final_scores=0.58193916,
final_local_scores=0.60834926, pem=0.0, similarity=22.244068]
8: 481609 Abimael_Guzmán
[pem_pos=0, final_scores=0.5095388,
final_local_scores=0.5673237, pem=0.0, similarity=21.019783]
9: 170691 Peru
[pem_pos=0, final_scores=0.64022607,
final_local_scores=0.6245356, pem=-0.23383401,
similarity=23.998728]

Sample: b'92_Former'
Former [1 1 Surinam 1] rebel leader held after shooting . [2 2 PARAMARIBO 2] , [3 3 Surinam 3] 1996-08-21 Flamboyant former [4 4 Surinamese 4] rebel leader Ronny Brunswijk was in custody on Wednesday charged with attempted murder , police said . Brunswijk turned himself into police after Freddy Pinas , a Surinamese-born visitor from the [5 5 Netherlands 5]

, accused Brunswijk of trying to kill him on Sunday after a bar-room brawl in the small mining town of [6 6 Moengo 6] , about 56 miles (90 km) east of [7 7 Paramaribo 7] , said police spokesman Ro Gajadhar . Pinas , showing cuts and bruises on his face , told reporters the former head of the feared Jungle Command had tried and failed to shoot him after Pinas objected to Brunswijk 's advances toward his wife . Pinas said Brunswijk then ordered his bodyguards to beat him up . Brunswijk , 35 , denied the charges and said he had merely defended himself when Pinas attacked him with a bottle . It was the second time Brunswijk had been charged with attempted murder in less than two years . In 1994 he served two months in prison for shooting a thief in the buttocks . Brunswijk led a rebel group of about 1,000 in a 1986 uprising against the regime of military strongman [8 8 De si Bouterse 8] . The conflict , which killed more than 500 and caused thousands to flee to neighbouring [9 9 French Guiana 9] in the late 1980s , eventually paved the way to democratic elections in 1991 . Despite numerous problems with authorities , Brunswijk went on to become a successful businessman with mining and logging interests . He also manages and occasionally plays for one of the leading local soccer teams .

1: 26828 Suriname
2: 57370 Paramaribo
3: 26828 Suriname
4: 26828 Suriname
5: 21148 Netherlands
6: 2418398 Moengo
7: 57370 Paramaribo
8: 105961 Dési_Bouterse
9: 21350970 French_Guiana
1: 26828 Suriname
[pem_pos=0, final_scores=0.574403,
final_local_scores=0.5305691, pem=0.0, similarity=19.922953]
2: 57370 Paramaribo
[pem_pos=0, final_scores=0.47360557,
final_local_scores=0.45622092, pem=0.0, similarity=17.704256]
3: 26828 Suriname
[pem_pos=0, final_scores=0.49563628,
final_local_scores=0.4445284, pem=0.0, similarity=17.355328]
4: 26828 Suriname
[pem_pos=0, final_scores=0.41578096,
final_local_scores=0.35729858, pem=0.0, similarity=14.752215]
5: 21148 Netherlands
[pem_pos=0, final_scores=0.5381339,
final_local_scores=0.54594445, pem=-0.24940795,
similarity=21.738104]
6: 2418398 Moengo
[pem_pos=0, final_scores=0.52725714,
final_local_scores=0.6847552, pem=0.0, similarity=24.524174]
7: 57370 Paramaribo

```
[pem_pos=0, final_scores=0.42748094,
final_local_scores=0.40583676, pem=0.0, similarity=16.200691]
8: 105961 Dési_Bouterse
    [pem_pos=0, final_scores=0.4808312,
final_local_scores=0.5716506, pem=0.0, similarity=21.148909]
9: 21350970 French_Guiana
    [pem_pos=0, final_scores=0.3549397,
final_local_scores=0.38526702, pem=0.0, similarity=15.586849]
```

Sample: b'93_Noisy'

Noisy saw leads [1 1 Thai 1] police to heroin hideaway . [2 2 BANGKOK 2] 1996-08-22 A [3 3 Hong Kong 3] carpenter was arrested in the [4 4 Thai 4] seaside town of [5 5 Pattaya 5] after police seized 18 kg (39.7 pounds) of heroin following complaints by residents of a noisy saw , police said on Thursday . Cheung Siu Man , 40 , was arrested late on Wednesday after police searched a house and found heroin in bags and hidden in hollow spaces in wooden planks , police said . The suspect said he was hired to make a wooden box from the planks in order to hide the heroin . Police went to the house after receiving complaints of sawing during the night over the course of several days . When they arrived to investigate , police saw people escaping from the back door so they decided to search the house . The seized heroin has an estimated street value of about 300 million baht (\$ 12 million) , police said . Officials are now hunting for the suspect 's collaborators , police said . Cheung was being detained pending formal charges , police said .

```
1: 30128 Thailand
2: 56615 Bangkok
3: 13404 Hong_Kong
4: 30128 Thailand
5: 712796 Pattaya
1: 30128 Thailand
    [pem_pos=2, final_scores=0.72665364,
final_local_scores=0.6968939, pem=-1.7504287,
similarity=34.405518]
2: 56615 Bangkok
    [pem_pos=0, final_scores=0.54368275,
final_local_scores=0.50368565, pem=-0.0005963434,
similarity=19.12394]
3: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.50807184,
final_local_scores=0.4841829, pem=-0.22402541,
similarity=19.75698]
4: 30128 Thailand
    [pem_pos=2, final_scores=0.6730086,
final_local_scores=0.6382949, pem=-1.7504287,
similarity=32.656807]
5: 712796 Pattaya
    [pem_pos=0, final_scores=0.3853398,
final_local_scores=0.37699717, pem=0.0, similarity=15.340061]
```

Sample: b'94_Australia'

[1 1 Australia 1] foreign minister arrives in [2 2 China 2] .
[3 3 BEIJING 3] 1996-08-22 [4 4 Australian 4] Foreign
Minister [5 5 Alexander Downer 5] arrived in [6 6 Beijing 6]
on Thursday for a four-day visit that follows rising friction
between the two nations in recent weeks . [7 7 Downer 7] was
to meet [8 8 Chinese 8] Foreign Minister [9 9 Qian Qichen 9]
and sign an agreement on an [10 10 Australian 10] consulate
in [11 11 Hong Kong 11] , an official of the [12 12 Australian
12] embassy in [13 13 Beijing 13] said . [14 14 China 14]
will resume sovereignty over [15 15 Hong Kong 15] , a [16 16 B
ritish 16] colony , in mid-1997 . Relations between [17 17 Ch
ina 17] and [18 18 Australia 18] have been strained in recent
weeks because of [19 19 Australia 19] 's plan to sell uranium
to [20 20 China 20] 's rival [21 21 Taiwan 21] . Other issues
affecting ties include plans by an [22 22 Australian 22]
cabinet minister to visit [23 23 Taiwan 23] , a security pact
between 1 Canberra and 2 Washington and a possible visit to [2
4 24 Australia 24] next month by [25 25 Tibet 25] 's exiled
spiritual leader the [26 26 Dalai Lama 26] . [27 27 Downer 27]
is the first [28 28 Australian 28] minister to visit [29 29 Ch
ina 29] since the new conservative government took office in [
30 30 Canberra 30] in March .

1: 4689264 Australia
2: 5405 China
3: 18603746 Beijing
4: 4689264 Australia
5: 331078 Alexander_Downer
6: 18603746 Beijing
7: 331078 Alexander_Downer
8: 5405 China
9: 728119 Qian_Qichen
10: 4689264 Australia
11: 13404 Hong_Kong
12: 4689264 Australia
13: 18603746 Beijing
14: 5405 China
15: 13404 Hong_Kong
16: 31717 United_Kingdom
17: 5405 China
18: 4689264 Australia
19: 4689264 Australia
20: 5405 China
21: 25734 Taiwan
22: 4689264 Australia
23: 25734 Taiwan
24: 4689264 Australia
25: 31516 Tibet
26: 20238168 14th_Dalai_Lama
27: 331078 Alexander_Downer
28: 4689264 Australia

```
29: 5405 China
30: 51983 Canberra
1: 4689264 Australia
    [winner: 51983 Canberra, pem_pos=0,
final_scores=-0.19456041, final_local_scores=-0.26845855,
pem=-0.091910355, similarity=-3.4218028|
     RECALL MISS: 4689264 Australia]
2: 3434750 United_States
    [winner: 108956 Washington, D.C., pem_pos=0,
final_scores=-0.1615655, final_local_scores=-0.20518902,
pem=-1.7496481, similarity=7.481321|
     RECALL MISS: 3434750 United_States]
1: 4689264 Australia
    [pem_pos=1, final_scores=0.66500586,
final_local_scores=0.6362502, pem=-0.17168477,
similarity=24.010338]
2: 5405 China
    [pem_pos=0, final_scores=0.54054326,
final_local_scores=0.5035717, pem=-0.15669827,
similarity=19.969446]
3: 18603746 Beijing
    [pem_pos=0, final_scores=0.49770838,
final_local_scores=0.4980747, pem=-0.042161036,
similarity=19.182535]
4: 4689264 Australia
    [pem_pos=0, final_scores=0.77011925,
final_local_scores=0.7510706, pem=-0.491009,
similarity=29.173342]
5: 331078 Alexander_Downer
    [pem_pos=0, final_scores=0.6725728,
final_local_scores=0.73943055, pem=0.0, similarity=26.155796]
6: 18603746 Beijing
    [pem_pos=0, final_scores=0.35986984,
final_local_scores=0.3475071, pem=-0.042161036,
similarity=14.689296]
7: 331078 Alexander_Downer
    [pem_pos=0, final_scores=0.9690953,
final_local_scores=1.071964, pem=0.0, similarity=36.079254]
8: 5405 China
    [pem_pos=1, final_scores=0.76662403,
final_local_scores=0.75053054, pem=-1.4196125,
similarity=34.207115]
9: 728119 Qian_Qichen
    [pem_pos=0, final_scores=0.5785656,
final_local_scores=0.720753, pem=0.0, similarity=25.598421]
10: 4689264 Australia
    [pem_pos=0, final_scores=0.6780359,
final_local_scores=0.65048355, pem=-0.491009,
similarity=26.171623]
11: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5286947,
final_local_scores=0.4967748, pem=-0.22402541,
```

```
similarity=20.13275]
12: 4689264 Australia
    [pem_pos=0, final_scores=0.52972573,
final_local_scores=0.48847723, pem=-0.491009,
similarity=21.33703]
13: 18603746 Beijing
    [pem_pos=0, final_scores=0.4380679,
final_local_scores=0.4329266, pem=-0.042161036,
similarity=17.238384]
14: 5405 China
    [pem_pos=0, final_scores=0.56190056,
final_local_scores=0.5269013, pem=-0.15669827,
similarity=20.66565]
15: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.4949785,
final_local_scores=0.45994496, pem=-0.22402541,
similarity=19.033674]
16: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.56431973,
final_local_scores=0.55214524, pem=-0.48936236,
similarity=23.228058]
17: 5405 China
    [pem_pos=0, final_scores=0.58090353,
final_local_scores=0.54765916, pem=-0.15669827,
similarity=21.285105]
18: 4689264 Australia
    [pem_pos=1, final_scores=0.60222155,
final_local_scores=0.56766796, pem=-0.17168477,
similarity=21.963707]
19: 4689264 Australia
    [pem_pos=1, final_scores=0.6568113,
final_local_scores=0.6272989, pem=-0.17168477,
similarity=23.743212]
20: 5405 China
    [pem_pos=0, final_scores=0.5520628,
final_local_scores=0.5161551, pem=-0.15669827,
similarity=20.344961]
21: 25734 Taiwan
    [pem_pos=0, final_scores=0.50225055,
final_local_scores=0.46344352, pem=-0.077886656,
similarity=18.343351]
22: 4689264 Australia
    [pem_pos=0, final_scores=0.66288465,
final_local_scores=0.6339331, pem=-0.491009,
similarity=25.677725]
23: 25734 Taiwan
    [pem_pos=0, final_scores=0.41506368,
final_local_scores=0.36820513, pem=-0.077886656,
similarity=15.501247]
24: 4689264 Australia
    [pem_pos=0, final_scores=0.6825939,
final_local_scores=0.65546244, pem=-0.17168477,
```

```
similarity=24.58367]
25: 31516 Tibet
    [pem_pos=0, final_scores=0.46722734,
final_local_scores=0.48119617, pem=-0.107068084,
similarity=19.031818]
26: 20238168 14th_Dalai_Lama
    [pem_pos=1, final_scores=0.1818767,
final_local_scores=0.206976, pem=-0.8232499,
similarity=14.743251]
27: 331078 Alexander_Downer
    [pem_pos=0, final_scores=0.91429764,
final_local_scores=1.012106, pem=0.0, similarity=34.292973]
28: 4689264 Australia
    [pem_pos=0, final_scores=0.69295925,
final_local_scores=0.66678506, pem=-0.491009,
similarity=26.658092]
29: 5405 China
    [pem_pos=0, final_scores=0.51080626,
final_local_scores=0.47108853, pem=-0.15669827,
similarity=19.000084]
30: 51983 Canberra
    [pem_pos=0, final_scores=0.455617,
final_local_scores=0.44678575, pem=-0.091910355,
similarity=17.922514]
```

Sample: b'95_Palestinians'

[1 1 Palestinians 1] accuse PA of banning books . [2 2 NABLUS 2] , West Bank 1996-08-22 A West Bank bookseller charged on Thursday that the Palestinian Information Ministry has forced him to sign an undertaking not to distribute books written by critics of Israeli-PLO self-rule deals . " They made me sign an undertaking not to sell the books to anyone at the risk of legal action . One official told me ' you have to either destroy the books or return them to [3 3 Amman 3] ' , " Daoud Makkawi , owner of the Nablus-based al-Risala bookshop , told [4 4 Reuters 4] . He said ministry officials made him sign this a few weeks ago after he brought about a dozen copies from [5 5 Jordan 5] of a book by [6 6 Edward Said 6] , a prominent scholar at [7 7 New York City 7] 's [8 8 Columbia University 8] . [9 9 Said 9] , a [10 10 U.S. 10] citizen of [11 11 Palestinian 11] origin , has been an outspoken critic of the 1993 Israeli-PLO self-rule deal and has written at least two books on the accord . On Wednesday a bookseller in the West Bank town of [12 12 Ramallah 12] said police about a month ago confiscated several copies of two of [13 13 Said 13] 's books on the Israel-PLO self-rule deals . Palestinian Information Ministry Director-General Mutawakel Taha denied that ministry officials forced anyone to sign any undertaking and insisted that the [14 14 Palestinian Authority 14] has no plans to censor books . " There is no strategy to ban books or to suppress freedom of expression in any form whatsoever , " Taha told [15 15 Reuters 15] . But Taha said that the

absence of relevant legislations may have resulted in some mistakes by some security officials . " This may explain some mistakes against some journalists and writers , " he said . Daoud said books by other authors , including [16 16 British 1 6] Journalist [17 17 Patrick Seale 17] , were also banned . He said that security officials often visit his shop to make sure he was not selling the books . " I think this is a bad beginning . If we have confidence , why should we be afraid of the other opinion ? " Daoud said . Thousands of books were banned from sale in the West Bank and [18 18 Gaza Strip 18] by the [19 19 Israeli 19] military authorities before the [20 20 Jewish 20] state handed over parts of the two areas to the PLO under a self-rule deal in 1994 .

1: 241405 State_of_Palestine
2: 233369 Nablus
3: 45007 Amman
4: 18998750 Reuters
5: 7515964 Jordan
6: 18247304 Edward_Said
7: 645042 New_York_City
8: 6310 Columbia_University
9: 18247304 Edward_Said
10: 3434750 United_States
11: 241405 State_of_Palestine
12: 44132 Ramallah
13: 18247304 Edward_Said
14: 24093 Palestinian_National_Authority
15: 18998750 Reuters
16: 31717 United_Kingdom
17: 7073056 Patrick_Seale
18: 12047 Gaza_Strip
19: 9282173 Israel
20: 25955086 Jews
1: 241405 State_of_Palestine
[pem_pos=5, final_scores=0.49215668,
final_local_scores=0.50603783, pem=-1.7909594,
similarity=28.930407]
2: 233369 Nablus
[pem_pos=0, final_scores=0.37697917,
final_local_scores=0.39127088, pem=0.0, similarity=15.766017]
3: 45007 Amman
[pem_pos=0, final_scores=0.39779228,
final_local_scores=0.38430506, pem=0.0, similarity=15.558142]
4: 18998750 Reuters
[pem_pos=0, final_scores=0.4290734,
final_local_scores=0.43230617, pem=0.0, similarity=16.990591]
5: 7515964 Jordan
[pem_pos=0, final_scores=0.5291926,
final_local_scores=0.5148953, pem=-0.53647494,
similarity=22.37265]
6: 18247304 Edward_Said
[pem_pos=0, final_scores=0.49378937,

```
final_local_scores=0.5057149, pem=0.0, similarity=19.181253]
7: 645042 New_York_City
    [pem_pos=0, final_scores=0.6101734,
final_local_scores=0.64897496, pem=0.0, similarity=23.45642]
8: 6310 Columbia_University
    [pem_pos=0, final_scores=0.38063663,
final_local_scores=0.4142571, pem=0.0, similarity=16.451971]
9: 18247304 Edward_Said
    [pem_pos=0, final_scores=0.63160634,
final_local_scores=0.6669855, pem=0.0, similarity=23.99389]
10: 3434750 United_States
    [pem_pos=0, final_scores=0.49141085,
final_local_scores=0.50430745, pem=-0.14075686,
similarity=19.90471]
11: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.6359876,
final_local_scores=0.66315126, pem=-2.8508017,
similarity=39.382572]
12: 44132 Ramallah
    [pem_pos=0, final_scores=0.41572744,
final_local_scores=0.38972414, pem=0.0, similarity=15.719858]
13: 18247304 Edward_Said
    [pem_pos=0, final_scores=0.62608826,
final_local_scores=0.6609579, pem=0.0, similarity=23.814014]
14: 24093 Palestinian_National_Authority
    [pem_pos=0, final_scores=0.40871906,
final_local_scores=0.39992017, pem=0.0, similarity=16.024128]
15: 18998750 Reuters
    [pem_pos=0, final_scores=0.36375922,
final_local_scores=0.3609604, pem=0.0, similarity=14.861491]
16: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.56185186,
final_local_scores=0.58025396, pem=-0.48936236,
similarity=24.066877]
17: 7073056 Patrick_Seale
    [pem_pos=0, final_scores=0.42335922,
final_local_scores=0.47404277, pem=0.0, similarity=18.236095]
18: 12047 Gaza_Strip
    [pem_pos=0, final_scores=0.40013486,
final_local_scores=0.39043486, pem=0.0, similarity=15.741068]
19: 9282173 Israel
    [pem_pos=0, final_scores=0.56731457,
final_local_scores=0.5429905, pem=-0.34790626,
similarity=22.1856]
20: 25955086 Jews
    [pem_pos=1, final_scores=0.36761218,
final_local_scores=0.36085558, pem=-0.80476934,
similarity=19.234823]
```

Sample: b'96_Egypt'

```
[1 1 Egypt 1] blames [2 2 Istanbul 2] control tower for
accident . [3 3 CAIRO 3] 1996-08-22 The chairman of national
```

carrier [4 4 EgyptAir 4] on Thursday blamed the control tower at [5 5 Istanbul 5] airport for the [6 6 EgyptAir 6] plane accident . Twenty people were injured on Wednesday when the [7 7 EgyptAir 7] [8 8 Boeing 8] 707 overshot the runway , caught fire , hit a taxi and skipped across a road onto a railway line . Chairman Mohamed Fahim Rayyan told a news conference at [9 9 Cairo 9] airport : " The control tower should have allocated the plane another runway , instead of the one the plane landed on . " " The one it landed on is 2,250 metres (2,460 yards) long while the other one if more than 3,000 metres (3,300 yards) long and is less steep , " he added . He said a Turkish civil aviation authority official had made the same point and he noted that a Turkish plane had a similar accident there in 1994 . The [10 10 EgyptAir 10] pilot blamed Turkish airport staff for misleading him . The landing took place after a rainstorm . " Its not an accident . It 's very wet . The brake action is very poor and the tower said it 's medium . That 's wrong , " the pilot told private Ihlas news agency in [11 11 English 11] .

1: 8087628 Egypt
2: 3391396 Istanbul
3: 6293 Cairo
4: 319357 EgyptAir
5: 3391396 Istanbul
6: 319357 EgyptAir
7: 319357 EgyptAir
8: 18933266 Boeing
9: 6293 Cairo
10: 319357 EgyptAir
11: 8569916 English_language
1: 8087628 Egypt
[pem_pos=0, final_scores=0.5461055,
final_local_scores=0.54002166, pem=-0.27480945,
similarity=21.699493]
2: 3391396 Istanbul
[pem_pos=0, final_scores=0.50586194,
final_local_scores=0.506782, pem=-0.025284406,
similarity=19.3506]
3: 6293 Cairo
[pem_pos=0, final_scores=0.62847596,
final_local_scores=0.61023533, pem=-0.15211283,
similarity=23.127567]
4: 319357 EgyptAir
[pem_pos=0, final_scores=0.5134689,
final_local_scores=0.4998753, pem=0.0, similarity=19.006989]
5: 3391396 Istanbul
[pem_pos=0, final_scores=0.390081,
final_local_scores=0.3803088, pem=-0.025284406,
similarity=15.576386]
6: 319357 EgyptAir
[pem_pos=0, final_scores=0.42128956,
final_local_scores=0.39918333, pem=0.0, similarity=16.00214]

```
7: 319357 EgyptAir
    [pem_pos=0, final_scores=0.45628434,
final_local_scores=0.43740982, pem=0.0, similarity=17.142895]
8: 18933266 Boeing
    [pem_pos=0, final_scores=0.32360023,
final_local_scores=0.39919674, pem=0.0, similarity=16.00254]
9: 6293 Cairo
    [pem_pos=0, final_scores=0.51278937,
final_local_scores=0.48386526, pem=-0.15211283,
similarity=19.356432]
10: 319357 EgyptAir
    [pem_pos=0, final_scores=0.4327256,
final_local_scores=0.41167545, pem=0.0, similarity=16.37493]
11: 8569916 English_language
    [pem_pos=0, final_scores=0.55545384,
final_local_scores=0.6379184, pem=-0.640266,
similarity=26.608335]
```

Sample: b'97_Egypt'

[1 1 Egypt 1] wants nothing to do with [2 2 Sudanese 2] rulers . [3 3 CAIRO 3] 1996-08-22 The [4 4 Egyptian 4] government will have nothing more to do with the [5 5 Sudanese 5] government because it continues to shelter and support [6 6 Egyptian 6] militants , President [7 7 Hosni Mubarak 7] said in a speech on Thursday . [8 8 Egypt 8] says the [9 9 Sudanese 9] government helped the [1 Moslem 1] militants who tried to kill [10 10 Mubarak 10] in Addis Ababa last year . It sponsored last week 's [11 11 U.N. Security Council 11] resolution threatening a ban on [12 12 Sudanese 12] flights abroad if [2 Khartoum 2] does not hand over three men accused in the Addis Ababa incident . The sanctions will come into effect in November if [13 13 Sudan 13] fails to extradite the men , but [14 14 Sudan 14] says it cannot hand them over to [1 5 15 Ethiopia 15] for trial because they are not in [16 16 Sudan 16] . " We are still eager that nothing should affect the [17 17 Sudanese 17] people but we will not deal with the current regime or the [18 18 Turabi 18] front or whatever , " [19 19 Mubarak 19] told a group of academics . [20 20 Hassan a l-Turabi 20] is the leader of the [21 21 National Islamic Front 21] , the political force behind the [22 22 Sudanese 22] government . " I do n't want to go into more details than that but there are more details and they are bitter . There are terrorists they are sheltering and they make [23 23 Sudanese 23] passorts for them and they get paid by them , " [24 24 Mubarak 24] said . He did not say if [25 25 Egypt 25] would go so far as to break relations , a step it has been reluctant to take , ostensibly because it would affect ordinary [26 26 Sudanese 26] .

```
1: 8087628 Egypt
2: 27421 Sudan
3: 6293 Cairo
4: 8087628 Egypt
```

5: 27421 Sudan
6: 8087628 Egypt
7: 51884 Hosni_Mubarak
8: 8087628 Egypt
9: 27421 Sudan
10: 51884 Hosni_Mubarak
11: 31956 United_Nations_Security_Council
12: 27421 Sudan
13: 27421 Sudan
14: 27421 Sudan
15: 187749 Ethiopia
16: 27421 Sudan
17: 27421 Sudan
18: 1088372 Hassan_al-Turabi
19: 51884 Hosni_Mubarak
20: 1088372 Hassan_al-Turabi
21: 1580709 National_Islamic_Front
22: 27421 Sudan
23: 27421 Sudan
24: 51884 Hosni_Mubarak
25: 8087628 Egypt
26: 27421 Sudan
1: 8087628 Egypt
[pem_pos=0, final_scores=0.5672254,
final_local_scores=0.52839875, pem=-0.27480945,
similarity=21.352642]
2: 27421 Sudan
[pem_pos=0, final_scores=0.5281866,
final_local_scores=0.4763118, pem=-0.19893567,
similarity=19.38565]
3: 6293 Cairo
[pem_pos=0, final_scores=0.59359735,
final_local_scores=0.58460355, pem=-0.15211283,
similarity=22.362663]
4: 8087628 Egypt
[pem_pos=0, final_scores=0.4468189,
final_local_scores=0.39687294, pem=-1.052751,
similarity=21.658216]
5: 27421 Sudan
[pem_pos=0, final_scores=0.5442745,
final_local_scores=0.4938854, pem=-0.19893567,
similarity=19.910082]
6: 8087628 Egypt
[pem_pos=0, final_scores=0.5212994,
final_local_scores=0.47823155, pem=-1.052751,
similarity=24.086119]
7: 51884 Hosni_Mubarak
[pem_pos=0, final_scores=0.5739631,
final_local_scores=0.566273, pem=0.0, similarity=20.988428]
8: 8087628 Egypt
[pem_pos=0, final_scores=0.58160067,
final_local_scores=0.54410154, pem=-0.27480945,

```
similarity=21.821245]
9: 27421 Sudan
    [pem_pos=0, final_scores=0.5198834,
final_local_scores=0.46724182, pem=-0.19893567,
similarity=19.114985]
10: 51884 Hosni_Mubarak
    [pem_pos=0, final_scores=0.53452104,
final_local_scores=0.528318, pem=0.0, similarity=19.855776]
11: 31956 United_Nations_Security_Council
    [pem_pos=0, final_scores=0.44040352,
final_local_scores=0.47888154, pem=0.0, similarity=18.380493]
12: 27421 Sudan
    [pem_pos=0, final_scores=0.55016387,
final_local_scores=0.50031865, pem=-0.19893567,
similarity=20.102062]
13: 27421 Sudan
    [pem_pos=0, final_scores=0.54656047,
final_local_scores=0.49638247, pem=-0.2848764,
similarity=20.451958]
14: 27421 Sudan
    [pem_pos=0, final_scores=0.59463364,
final_local_scores=0.5488952, pem=-0.2848764,
similarity=22.019043]
15: 187749 Ethiopia
    [pem_pos=0, final_scores=0.60584813,
final_local_scores=0.60178626, pem=-0.23532115,
similarity=23.327929]
16: 27421 Sudan
    [pem_pos=0, final_scores=0.52670866,
final_local_scores=0.47469735, pem=-0.2848764,
similarity=19.80483]
17: 27421 Sudan
    [pem_pos=0, final_scores=0.52129644,
final_local_scores=0.4687854, pem=-0.19893567,
similarity=19.161047]
18: 1088372 Hassan_al-Turabi
    [pem_pos=0, final_scores=0.43088,
final_local_scores=0.4554497, pem=0.0, similarity=17.68124]
19: 51884 Hosni_Mubarak
    [pem_pos=0, final_scores=0.5007993,
final_local_scores=0.49148208, pem=0.0, similarity=18.75652]
20: 1088372 Hassan_al-Turabi
    [pem_pos=0, final_scores=0.63284355,
final_local_scores=0.6760642, pem=0.0, similarity=24.264816]
21: 1580709 National_Islamic_Front
    [pem_pos=0, final_scores=0.5316069,
final_local_scores=0.55813634, pem=0.0, similarity=20.745615]
22: 27421 Sudan
    [pem_pos=0, final_scores=0.5282354,
final_local_scores=0.4763651, pem=-0.19893567,
similarity=19.387241]
23: 27421 Sudan
```

```
[pem_pos=0, final_scores=0.50347406,
final_local_scores=0.4493171, pem=-0.19893567,
similarity=18.580074]
24: 51884 Hosni_Mubarak
    [pem_pos=0, final_scores=0.48652714,
final_local_scores=0.4758919, pem=0.0, similarity=18.291277]
25: 8087628 Egypt
    [pem_pos=0, final_scores=0.55713445,
final_local_scores=0.51737595, pem=-0.27480945,
similarity=21.0237]
26: 27421 Sudan
    [pem_pos=0, final_scores=0.43125647,
final_local_scores=0.37043035, pem=-0.19893567,
similarity=16.225937]
1: 6037917 Islam
    [pem_pos=1, final_scores=0.18603088,
final_local_scores=0.1878488, pem=-1.2150909,
similarity=16.303349]
2: 17320 Khartoum
    [pem_pos=0, final_scores=0.04824333,
final_local_scores=-0.013340235, pem=-0.03966028,
similarity=3.9072926]
```

Sample: b'98_Turkish'

Turkish shares shed gains in profit-taking . [1 1 ISTANBUL 1] 1996-08-22 Turkish shares ended lower on Thursday , shedding gains of earlier in the week amid profit-taking sales , brokers said . The IMKB-100 lost 0.19 percent or 123.89 points to end at 64,178.78 . Gains so far this week have totalled 2.92 percent . Daily volume dropped to 7.2 trillion lira from Wednesday 's 7.8 trillion lira . " Profit-taking sales in the afternoon showed the latest gains of the index were actually a reaction rise . I expect the market to go as far down as 63,000 tomorrow if sales continue , " said Burcin Mavituna from Interbank . Brokers said profit taking sales had come especially as the index approached the 65,000 resistance level . They said the index could also rise towards 65,000 if the cheap share prices attracted buyers . The market had its first resistance at 67,000 if it pierced 65,000 , they added . The session 's most active shares were those of Isbank gained 300 lira to 8,600 . Shares of utility Cukurova lost 3,000 lira to 67,000 . The 85-share industrial index lost 0.47 percent to 70,848.86 and the 15-share financial index rose by 0.55 percent to 55,929.89 . Of the 218 shares traded , gainers outdid losers by 100 to 64 and 54 shares were stable . -- Istanbul Newsroom , +90-212-275 0875 SA

1: 3391396 Istanbul

1: 3391396 Istanbul

```
[pem_pos=0, final_scores=0.20377986,
final_local_scores=0.49518573, pem=-0.025284406,
similarity=19.004543]
```

Sample: b'99_Miss'

[1 1 Miss Universe 1] hides behind veil of silence . Kieran Murray [2 2 LAS CRUCES 2] , N.M. 1996-08-22 [3 3 Miss Universe 3] , [4 4 Venezuela 4] 's [5 5 Alicia Machado 5] , left [6 6 New Mexico 6] on Thursday , refusing to answer questions about her weight or claims she was told to either go on a crash diet or give up her title . [7 7 Machado 7] , 19 , flew to [8 8 Los Angeles 8] after slipping away from the [9 9 New Mexico 9] desert town of [10 10 Las Cruces 10] , where she attended the 1996 Miss Teen USA pageant on Wednesday . While [11 11 Machado 11] was not a contestant here , she came under intense scrutiny following reports she was given an ultimatum by [12 1 2 Los Angeles-based 12] Miss Universe Inc. to drop 27 pounds (12 kg) in two weeks or risk losing her crown . In [13 13 Venezuela 13] , her mother told [14 14 Reuters 14] that [15 15 Machado 15] had a swollen face when she left home two weeks ago because she had her wisdom teeth extracted . Marta Fajardo insisted her daughter , who weighed 112 pounds (51 kg) when she won the [16 16 Miss Universe 16] title in [17 17 Las Vegas 17] in May , had perfectly normal eating habits . " Everybody has their own addiction to something or other but it 's not as if she eats cakes like crazy , " she said . Organisers flatly denied ever threatening [18 18 Machado 18] but immediately put her under wraps and blocked access to her . Dressed in a black strapless evening gown at Wednesday 's pageant , [19 19 Machado 19] was clearly heavier than the contestants but still won rave reviews after her brief appearance on stage . " Are you kidding ? She 's fantastic , " said Nikki Campbell , 28 , who went to the pageant . " She looked great . Very sexy . " [20 20 Machado 20] 's publicists said on Thursday she was scheduled to stay in [21 21 Los Angeles 21] for promotional work with sponsors before returning to [22 22 Venezuela 22] on Sept . 5 . Beauty queens are high-profile personalities in [23 23 Venezuela 23] and [24 24 Machado 24] 's alleged weight problem made front page news this week . It was an official of the Miss Venezuela Organisation who first said [25 25 Machado 25] had been told to lose weight fast . People close to her said she then eased up on her diet and indulged her passion for pasta and cake , but it was not clear how many pounds she gained and most people who saw her said she was still a long way from being fat . Martin Brooks , president of Miss Universe Inc , said he spoke with [26 26 Machado 26] to assure her that organisers were not putting pressure on her . " She 's fine with it . She wished , as we all did , that it had n't happened but she 's spiritually and mentally terrific . There 's no problem whatsoever , " he told [27 27 Reuters 27] . He said the lifestyle associated with being [28 28 Miss Universe 28] could make routine exercise difficult . " The problem is they travel so much and are so busy that the ability to have any type of regimented routine workout does n't exist . I dont

know if [29 29 Alicia 29] is working out . We have n't talked about it because it has n't been an issue , " he said .

1: 150340 Miss_Universe
2: 125874 Las_Cruces,_New_Mexico
3: 150340 Miss_Universe
4: 32374 Venezuela
5: 693474 Alicia_Machado
6: 21649 New_Mexico
7: 693474 Alicia_Machado
8: 18110 Los_Angeles
9: 21649 New_Mexico
10: 125874 Las_Cruces,_New_Mexico
11: 693474 Alicia_Machado
12: 18110 Los_Angeles
13: 32374 Venezuela
14: 18998750 Reuters
15: 693474 Alicia_Machado
16: 150340 Miss_Universe
17: 47737 Las_Vegas
18: 693474 Alicia_Machado
19: 693474 Alicia_Machado
20: 693474 Alicia_Machado
21: 18110 Los_Angeles
22: 32374 Venezuela
23: 32374 Venezuela
24: 693474 Alicia_Machado
25: 693474 Alicia_Machado
26: 693474 Alicia_Machado
27: 18998750 Reuters
28: 150340 Miss_Universe
29: 693474 Alicia_Machado
1: 150340 Miss_Universe
 [pem_pos=0, final_scores=0.47893244,
final_local_scores=0.471242, pem=-0.08560056,
similarity=18.618023]
2: 125874 Las_Cruces,_New_Mexico
 [pem_pos=0, final_scores=0.35884613,
final_local_scores=0.39571238, pem=0.0, similarity=15.89856]
3: 150340 Miss_Universe
 [pem_pos=0, final_scores=0.5280371,
final_local_scores=0.5248813, pem=-0.08560056,
similarity=20.218727]
4: 32374 Venezuela
 [pem_pos=0, final_scores=0.6265942,
final_local_scores=0.65818346, pem=-0.21226841,
similarity=24.885569]
5: 693474 Alicia_Machado
 [pem_pos=0, final_scores=0.57770747,
final_local_scores=0.56003946, pem=0.0, similarity=20.802408]
6: 21649 New_Mexico
 [pem_pos=0, final_scores=0.4792053,
final_local_scores=0.53506696, pem=-0.20017311,

```
similarity=21.14575]
7: 693474 Alicia_Machado
    [pem_pos=0, final_scores=0.5737873,
final_local_scores=0.5557572, pem=0.0, similarity=20.674618]
8: 18110 Los_Angeles
    [pem_pos=0, final_scores=0.52519816,
final_local_scores=0.5402364, pem=-0.100652285,
similarity=20.758808]
9: 21649 New_Mexico
    [pem_pos=0, final_scores=0.531773,
final_local_scores=0.5924892, pem=-0.20017311,
similarity=22.859344]
10: 125874 Las_Cruces,_New_Mexico
    [pem_pos=0, final_scores=0.39619607,
final_local_scores=0.43651152, pem=0.0, similarity=17.116087]
11: 693474 Alicia_Machado
    [pem_pos=0, final_scores=0.6029311,
final_local_scores=0.5875925, pem=0.0, similarity=21.624645]
12: 18110 Los_Angeles
    [pem_pos=0, final_scores=0.5675582,
final_local_scores=0.5865084, pem=0.0, similarity=21.592293]
13: 32374 Venezuela
    [pem_pos=0, final_scores=0.54034054,
final_local_scores=0.5639644, pem=-0.21226841,
similarity=22.073885]
14: 18998750 Reuters
    [pem_pos=0, final_scores=0.4034288,
final_local_scores=0.4740919, pem=0.0, similarity=18.23756]
15: 693474 Alicia_Machado
    [pem_pos=0, final_scores=0.52515835,
final_local_scores=0.50263757, pem=0.0, similarity=19.08942]
16: 150340 Miss_Universe
    [pem_pos=0, final_scores=0.5607653,
final_local_scores=0.56063193, pem=-0.08560056,
similarity=21.285597]
17: 47737 Las_Vegas
    [pem_pos=0, final_scores=0.37408024,
final_local_scores=0.37136883, pem=-0.49559593,
similarity=17.867228]
18: 693474 Alicia_Machado
    [pem_pos=0, final_scores=0.5317541,
final_local_scores=0.5098424, pem=0.0, similarity=19.304426]
19: 693474 Alicia_Machado
    [pem_pos=0, final_scores=0.5361503,
final_local_scores=0.5146446, pem=0.0, similarity=19.447735]
20: 693474 Alicia_Machado
    [pem_pos=0, final_scores=0.54795635,
final_local_scores=0.5275409, pem=0.0, similarity=19.832586]
21: 18110 Los_Angeles
    [pem_pos=0, final_scores=0.6268614,
final_local_scores=0.651288, pem=-0.100652285,
similarity=24.072807]
```

```
22: 32374 Venezuela
    [pem_pos=0, final_scores=0.51962894,
final_local_scores=0.5413401, pem=-0.21226841,
similarity=21.398731]
23: 32374 Venezuela
    [pem_pos=0, final_scores=0.52989805,
final_local_scores=0.5525576, pem=-0.21226841,
similarity=21.733482]
24: 693474 Alicia_Machado
    [pem_pos=0, final_scores=0.5350349,
final_local_scores=0.5134262, pem=0.0, similarity=19.411375]
25: 693474 Alicia_Machado
    [pem_pos=0, final_scores=0.5427204,
final_local_scores=0.52182144, pem=0.0, similarity=19.661905]
26: 693474 Alicia_Machado
    [pem_pos=0, final_scores=0.53092784,
final_local_scores=0.50893986, pem=0.0, similarity=19.277494]
27: 18998750 Reuters
    [pem_pos=0, final_scores=0.38015604,
final_local_scores=0.44866997, pem=0.0, similarity=17.47892]
28: 150340 Miss_Universe
    [pem_pos=0, final_scores=0.42735368,
final_local_scores=0.4149, pem=-0.08560056,
similarity=16.936666]
29: 693474 Alicia_Machado
    [pem_pos=0, final_scores=0.61039597,
final_local_scores=0.5957467, pem=0.0, similarity=21.867983]
```

Sample: b'100_Kevorkian'

[1 1 Kevorkian 1] attends third suicide in week . [2 2 PONTI AC 2] , Mich . 1996-08-22 Dr. [3 3 Jack Kevorkian 3] attended his third suicide in less than a week on Thursday , bringing the body of a 40-year-old [4 4 Missouri 4] woman suffering from multiple sclerosis to a hospital emergency room , doctors said . Dr Robert Aranosian , emergency room director at Pontiac Osteopathic Hospital , said [5 5 Kevorkian 5] brought in the body of Patricia Smith , of [6 6 Lees Summit 6] , Mo . , at midday and told doctors that she had been paralysed by the disease . It was his second assisted-suicide in 36 hours and the 37th that he has acknowledged attending since starting his crusade for doctor assisted suicide in 1990 . [7 7 Kevorkian 7] 's lawyer , [8 8 Geoffrey Fieger 8] , said those attending Smith 's death included her husband , David , a police officer , her father , James Poland , and [9 9 Kevorkian 9] . It was the first known time that a police officer has been president at the suicide of one of [10 10 Kevorkian 10] 's patients . He offered no details about the cause of Smith 's death or the location . She was a nurse who had " rapidly progressing multiple sclerosis . " On Tuesday night , [11 11 Kevorkian 11] attended the death of Louise Siebens , a 76-year-old [12 12 Texas 12] woman with amyotrophic lateral sclerosis , or [13 13 L

ou Gehrig 13] 's disease . On August 15 , [14 14 Kevorkian 14] helped Judith Curren , a 42-year-old [15 15 Massachusetts 15] nurse , who suffered from chronic fatigue syndrome , a non-terminal illness , to end her life .

1: 202835 Jack_Kevorkian
2: 118786 Pontiac,_Michigan
3: 202835 Jack_Kevorkian
4: 19571 Missouri
5: 202835 Jack_Kevorkian
6: 150725 Lee's_Summit,_Missouri
7: 202835 Jack_Kevorkian
8: 1957839 Geoffrey_Fieger
9: 202835 Jack_Kevorkian
10: 202835 Jack_Kevorkian
11: 202835 Jack_Kevorkian
12: 29810 Texas
13: 19381762 Lou_Gehrig
14: 202835 Jack_Kevorkian
15: 1645518 Massachusetts
1: 202835 Jack_Kevorkian
 [pem_pos=0, final_scores=0.6112265,
final_local_scores=0.57368994, pem=0.0, similarity=21.209764]
2: 118786 Pontiac,_Michigan
 [pem_pos=1, final_scores=0.44158238,
final_local_scores=0.49099332, pem=-1.6390301,
similarity=27.655233]
3: 202835 Jack_Kevorkian
 [pem_pos=0, final_scores=0.5482208,
final_local_scores=0.5048658, pem=0.0, similarity=19.155916]
4: 19571 Missouri
 [pem_pos=0, final_scores=0.42015022,
final_local_scores=0.4580376, pem=-0.27068004,
similarity=19.230469]
5: 202835 Jack_Kevorkian
 [pem_pos=0, final_scores=0.5291192,
final_local_scores=0.4840002, pem=0.0, similarity=18.533245]
6: 150725 Lee's_Summit,_Missouri
 [pem_pos=0, final_scores=0.46718264,
final_local_scores=0.5598255, pem=0.0, similarity=20.796022]
7: 202835 Jack_Kevorkian
 [pem_pos=0, final_scores=0.52844983,
final_local_scores=0.48326904, pem=0.0, similarity=18.511425]
8: 1957839 Geoffrey_Fieger
 [pem_pos=0, final_scores=0.513935,
final_local_scores=0.55591106, pem=0.0, similarity=20.679209]
9: 202835 Jack_Kevorkian
 [pem_pos=0, final_scores=0.53727376,
final_local_scores=0.49290782, pem=0.0, similarity=18.799065]
10: 202835 Jack_Kevorkian
 [pem_pos=0, final_scores=0.6182998,
final_local_scores=0.5814164, pem=0.0, similarity=21.440338]
11: 202835 Jack_Kevorkian

```
[pem_pos=0, final_scores=0.58940625,
final_local_scores=0.54985464, pem=0.0, similarity=20.498472]
12: 29810 Texas
    [pem_pos=0, final_scores=0.7538081,
final_local_scores=0.84023225, pem=-0.16191302,
similarity=30.044426]
13: 19381762 Lou_Gehrig
    [pem_pos=0, final_scores=0.35997885,
final_local_scores=0.46494377, pem=0.0, similarity=17.964561]
14: 202835 Jack_Kevorkian
    [pem_pos=0, final_scores=0.5885027,
final_local_scores=0.54886764, pem=0.0, similarity=20.469019]
15: 1645518 Massachusetts
    [pem_pos=0, final_scores=0.5459251,
final_local_scores=0.59991705, pem=-0.1423749,
similarity=22.766691]
```

Sample: b'101_Fairview'

```
[1 1 Fairview 1] , [2 2 Texas 2] , $ 1.82 million deal Baal
- Moody 's . [3 3 NEW YORK 3] 1996-08-22 Moody 's Investors
Service - Rating Announcement As of 08/21/96 . Issuer :
Fairview Town State : TX Rating : Baal Sale Amount :
1,820,000 Expected Sale Date : 08/27/96 -- U.S. Municipal
Desk , 212-859-1650
1: 135693 Fairview,_Texas
2: 29810 Texas
3: 645042 New_York_City
1: 135693 Fairview,_Texas
    [pem_pos=11, final_scores=0.35851753,
final_local_scores=0.46111262, pem=-3.3622322,
similarity=36.134575]
2: 29810 Texas
    [pem_pos=0, final_scores=0.71887803,
final_local_scores=0.7540779, pem=-0.16191302,
similarity=27.47341]
3: 645042 New_York_City
    [pem_pos=1, final_scores=0.39242786,
final_local_scores=0.44045615, pem=-1.7906687,
similarity=26.971737]
```

Sample: b'102_Defiant'

```
Defiant [1 1 U.S. 1] neo-Nazi jailed by [2 2 German 2]
court . Andrew Gray [3 3 HAMBURG 3] , [4 4 Germany 4]
1996-08-22 A [5 5 Hamburg 5] court sentenced [6 6 U.S. 6] [7 7
neo-Nazi 7] leader Gary Lauck on Thursday to four years in
prison for pumping banned extremist propaganda into [8 8 Germ
any 8] from his base in the [9 9 United States 9] . Lauck ,
from [10 10 Lincoln 10] , [11 11 Nebraska 11] , yelled a
 tirade of abuse at the court after his conviction for
inciting racial hatred . " The struggle will go on , " the 43-
year-old shouted in [12 12 German 12] before being escorted
out by security guards . Lauck 's lawyer vowed he would
```

appeal against the court 's decision , arguing that his client should have been set free because he had not committed any offence under [13 13 German 13] law . The [14 14 German 14] government hailed the conviction as a major victory in the fight against [15 15 neo-Nazism 15] . Lauck 's worldwide network has been the main source of anti-Semitic propaganda material flowing into [16 16 Germany 16] since the 1970s . " Lauck possessed a well-oiled propaganda machine , honed during more than 20 years , " presiding judge Guenter Bertram told the court . " He set up a propaganda cannon and fired it at [17 17 Germany 17] . " said Bertram , who also read out extracts from Lauck 's material praising [18 18 Hitler 18] as " the greatest of all leaders " and describing the [19 19 Nazi 19] slaughter of millions of [20 20 Jews 20] as a myth . Eager to put Lauck behind bars quickly and avoid a long and complex trial , prosecutor Bernd Mauruschat limited his charges to offences since 1994 . He had demanded a five-year jail term but said he was satisfied with the court 's sentence . Publishing and distributing [21 21 neo-Nazi 21] material is illegal in [22 22 Germany 22] but Lauck 's defence team had argued that [23 23 U.S 23] freedom of speech laws meant he was free to produce his swastika-covered books , magazines , videos and flags in his homeland . Interior Minister [24 24 Manfred Kanther 24] said in a statement he " welcomed the prosecution and conviction of one of the ringleaders of international [25 25 neo-Nazism 25] and biggest distributers of vicious racist publications " . " It is high time he was behind bars , " the opposition [26 26 Social Democrats 26] said in a statement . Lauck , dressed in a sober blue suit and sporting his trademark Hitleresque black moustache , showed no sign of emotion as Bertram spent more than an hour reading out the verdict and explaining the court 's decision . But as Lauck was about to be led away , he turned to reporters and blurted out a virtually incomprehensible quick-fire diatribe against the court . " Neither the [27 27 National Socialists 27] ([28 28 Nazis 28]) nor the communists dared to kidnap an [29 29 American 29] citizen , " he shouted , in an oblique reference to his extradition to [30 30 Germany 30] from [31 31 Denmark 31] . " That 's the truth . " His attorney , Hans-Otto Sieg , told reporters outside the courtroom that the judges had not explained how a [32 32 German 32] court could judge someone for actions carried out in the [33 33 United States 33] . Bertram said Lauck was obsessed by [34 34 Nazism 34] and devoted his life to leading his National Socialist German Workers ' Party Foreign Organisation (NSDAP-AO) , which derives its name from the full [35 35 German 35] title of [36 36 Hitler 36] 's [37 37 Nazi 37] party . During the three-month trial , the court dealt mainly with issues of the [38 38 NSDAP-AO 38] 's " NS Kampfruf " (" National Socialist Battle Cry ") magazine , filled with references to [39 39 Aryan 39] supremacy and defamatory statements about [40 40 Jews 40] .

The court rejected Sieg 's argument that Lauck 's extradition from [41 41 Denmark 41] , where he was arrested in March last year at the request of [42 42 German 42] authorities , was illegal . Lauck was also convicted of disseminating the symbols of anti-constitutional organisations . He will probably be free in around two and a half years . The court ruled that the 15 months he has spent in custody since his arrest should be subtracted from his prison term .

1: 3434750 United_States
2: 11867 Germany
3: 13467 Hamburg
4: 11867 Germany
5: 13467 Hamburg
6: 3434750 United_States
7: 54361 Neo-Nazism
8: 11867 Germany
9: 3434750 United_States
10: 17653 Lincoln,_Nebraska
11: 21647 Nebraska
12: 11884 German_language
13: 11867 Germany
14: 11867 Germany
15: 54361 Neo-Nazism
16: 11867 Germany
17: 11867 Germany
18: 2731583 Adolf_Hitler
19: 21212 Nazi_Germany
20: 25955086 Jews
21: 54361 Neo-Nazism
22: 11867 Germany
23: 3434750 United_States
24: 8517538 Manfred_Kanther
25: 54361 Neo-Nazism
26: 60585 Social_Democratic_Party_of_Germany
27: 21736 Nazi_Party
28: 21736 Nazi_Party
29: 3434750 United_States
30: 11867 Germany
31: 76972 Denmark
32: 11867 Germany
33: 3434750 United_States
34: 31045316 Nazism
35: 11884 German_language
36: 2731583 Adolf_Hitler
37: 21212 Nazi_Germany
38: 2570087 NSDAP/A0
39: 42277 Aryan_race
40: 25955086 Jews
41: 76972 Denmark
42: 11867 Germany
1: 3434750 United_States
[pem_pos=0, final_scores=0.45494252,

```
final_local_scores=0.435062, pem=-0.14075686,
similarity=17.83829]
2: 11867 Germany
    [pem_pos=2, final_scores=0.9078185,
final_local_scores=0.88696235, pem=-1.0465823,
similarity=36.249916]
3: 13467 Hamburg
    [pem_pos=0, final_scores=0.49139047,
final_local_scores=0.46526414, pem=-0.27202022,
similarity=19.453411]
4: 11867 Germany
    [pem_pos=0, final_scores=0.87719935,
final_local_scores=0.85236865, pem=-0.15529114,
similarity=30.37059]
5: 13467 Hamburg
    [pem_pos=0, final_scores=0.52764124,
final_local_scores=0.5048626, pem=-0.27202022,
similarity=20.635109]
6: 3434750 United_States
    [pem_pos=0, final_scores=0.47965366,
final_local_scores=0.46205515, pem=-0.14075686,
similarity=18.643818]
7: 54361 Neo-Nazism
    [pem_pos=0, final_scores=0.45013702,
final_local_scores=0.46506494, pem=0.0, similarity=17.968178]
8: 11867 Germany
    [pem_pos=0, final_scores=0.8621476,
final_local_scores=0.83592695, pem=-0.15529114,
similarity=29.879938]
9: 3434750 United_States
    [pem_pos=0, final_scores=0.5544967,
final_local_scores=0.54380983, pem=-0.073457584,
similarity=20.717558]
10: 17653 Lincoln,_Nebraska
    [pem_pos=3, final_scores=0.26848215,
final_local_scores=0.29436916, pem=-2.526223,
similarity=26.612268]
11: 21647 Nebraska
    [pem_pos=0, final_scores=0.50490993,
final_local_scores=0.54627454, pem=-0.4226844,
similarity=22.690258]
12: 11884 German_language
    [pem_pos=1, final_scores=0.37790143,
final_local_scores=0.3614642, pem=-0.981596,
similarity=20.214598]
13: 11867 Germany
    [pem_pos=2, final_scores=0.9375405,
final_local_scores=0.919429, pem=-1.0465823,
similarity=37.21879]
14: 11867 Germany
    [pem_pos=2, final_scores=0.882734,
final_local_scores=0.85956126, pem=-1.0465823,
```

```
similarity=35.432213]
15: 54361 Neo-Nazism
    [pem_pos=0, final_scores=0.4712416,
final_local_scores=0.48811847, pem=0.0, similarity=18.656141]
16: 11867 Germany
    [pem_pos=0, final_scores=0.79316,
final_local_scores=0.7605685, pem=-0.15529114,
similarity=27.631092]
17: 11867 Germany
    [pem_pos=0, final_scores=0.67836803,
final_local_scores=0.63517576, pem=-0.15529114,
similarity=23.889122]
18: 2731583 Adolf_Hitler
    [pem_pos=0, final_scores=0.44637746,
final_local_scores=0.43642265, pem=0.0, similarity=17.113436]
19: 21212 Nazi_Germany
    [pem_pos=1, final_scores=0.16138624,
final_local_scores=0.11565569, pem=-1.9176292,
similarity=17.969475]
20: 25955086 Jews
    [pem_pos=0, final_scores=0.5231892,
final_local_scores=0.523499, pem=-0.1252217,
similarity=20.392942]
21: 54361 Neo-Nazism
    [pem_pos=0, final_scores=0.3942575,
final_local_scores=0.40402508, pem=0.0, similarity=16.146627]
22: 11867 Germany
    [pem_pos=0, final_scores=0.68796736,
final_local_scores=0.6456615, pem=-0.15529114,
similarity=24.202038]
23: 3434750 United_States
    [pem_pos=0, final_scores=0.4636894,
final_local_scores=0.44461668, pem=0.0, similarity=17.357962]
24: 8517538 Manfred_Kanther
    [pem_pos=0, final_scores=0.44252986,
final_local_scores=0.59362674, pem=0.0, similarity=21.804718]
25: 54361 Neo-Nazism
    [pem_pos=0, final_scores=0.5122196,
final_local_scores=0.5328808, pem=0.0, similarity=19.991938]
26: 60585 Social_Democratic_Party_of_Germany
    [pem_pos=2, final_scores=0.21007769,
final_local_scores=0.20184222, pem=-1.541773,
similarity=18.497488]
27: 21736 Nazi_Party
    [pem_pos=1, final_scores=0.25911665,
final_local_scores=0.22135222, pem=-0.24943839,
similarity=12.051789]
28: 21736 Nazi_Party
    [pem_pos=2, final_scores=0.090561405,
final_local_scores=0.037231237, pem=-1.4827712,
similarity=13.264308]
29: 3434750 United_States
```

```
[pem_pos=0, final_scores=0.54477054,
final_local_scores=0.5331855, pem=-0.12994486,
similarity=20.707693]
30: 11867 Germany
    [pem_pos=0, final_scores=0.8665016,
final_local_scores=0.84068304, pem=-0.15529114,
similarity=30.02187]
31: 76972 Denmark
    [pem_pos=0, final_scores=0.5539444,
final_local_scores=0.53554565, pem=-0.26762357,
similarity=21.526842]
32: 11867 Germany
    [pem_pos=1, final_scores=0.8879611,
final_local_scores=0.86412424, pem=-1.0465823,
similarity=35.568382]
33: 3434750 United_States
    [pem_pos=0, final_scores=0.4892165,
final_local_scores=0.47250116, pem=-0.073457584,
similarity=18.589563]
34: 31045316 Nazism
    [pem_pos=0, final_scores=0.293656,
final_local_scores=0.2571201, pem=0.0, similarity=11.76269]
35: 11884 German_language
    [pem_pos=1, final_scores=0.29769313,
final_local_scores=0.27384883, pem=-0.981596,
similarity=17.59998]
36: 2731583 Adolf_Hitler
    [pem_pos=0, final_scores=0.40150958,
final_local_scores=0.38741136, pem=0.0, similarity=15.65084]
37: 21212 Nazi_Germany
    [pem_pos=1, final_scores=0.16554879,
final_local_scores=0.12020263, pem=-1.9176292,
similarity=18.105165]
38: 2570087 NSDAP/A0
    [pem_pos=0, final_scores=0.3097691,
final_local_scores=0.3666013, pem=0.0, similarity=15.029827]
39: 42277 Aryan_race
    [pem_pos=1, final_scores=0.187721,
final_local_scores=0.21380934, pem=-1.460011,
similarity=18.409977]
40: 25955086 Jews
    [pem_pos=1, final_scores=0.4596789,
final_local_scores=0.45412368, pem=-0.1252217,
similarity=18.322643]
41: 76972 Denmark
    [pem_pos=0, final_scores=0.5103716,
final_local_scores=0.487949, pem=-0.26762357,
similarity=20.106462]
42: 11867 Germany
    [pem_pos=2, final_scores=0.91637546,
final_local_scores=0.89630955, pem=-1.0465823,
similarity=36.528854]
```

Sample: b'103_UN'

[1 UN 1] official says [1 1 Iraqi 1] deal will occur " soon " . [2 2 UNITED NATIONS 2] 1996-08-22 A senior [3 3 U.N. 3] official said on Thursday he expected arrangements to implement the [4 4 Iraqi 4] oil-for-food deal could be completed " quite soon . " " I am reluctant to speculate but we are doing the preparations and the secretary-general is anxious to start the program , " said Undersecretary-General [5 5 Yasushi Akashi 5] . " It might be sooner than you think , " he told reporters after briefing the [6 6 Security Council 6] on arrangements for monitors needed to carry out the agreement . [7 7 Akashi 7] is head of the Department of Humanitarian affairs . His deputy earlier speculated at least 10 days .

1: 7515928 Iraq
2: 31769 United_Nations
3: 31769 United_Nations
4: 7515928 Iraq
5: 1271160 Yasushi_Akashi
6: 31956 United_Nations_Security_Council
7: 1271160 Yasushi_Akashi
1: 7515928 Iraq
 [pem_pos=0, final_scores=0.4557879,
final_local_scores=0.47150135, pem=-0.48469657,
similarity=20.796108]
2: 31769 United_Nations
 [pem_pos=0, final_scores=0.41169184,
final_local_scores=0.38009483, pem=0.0, similarity=15.4325]
3: 31769 United_Nations
 [pem_pos=0, final_scores=0.5053652,
final_local_scores=0.48241878, pem=0.0, similarity=18.486052]
4: 7515928 Iraq
 [pem_pos=0, final_scores=0.41355735,
final_local_scores=0.42537087, pem=-0.48469657,
similarity=19.419483]
5: 1271160 Yasushi_Akashi
 [pem_pos=0, final_scores=0.8044892,
final_local_scores=0.83724475, pem=0.0, similarity=29.074766]
6: 31956 United_Nations_Security_Council
 [pem_pos=0, final_scores=0.31360668,
final_local_scores=0.2957429, pem=0.0, similarity=12.915271]
7: 1271160 Yasushi_Akashi
 [pem_pos=0, final_scores=0.8327419,
final_local_scores=0.90004176, pem=0.0, similarity=30.948753]
1: 31769 United_Nations
 [pem_pos=0, final_scores=0.04768984,
final_local_scores=-0.032544933, pem=-0.06880474,
similarity=3.4926782]

Sample: b'104_Suspected'

Suspected killers of bishop dead -- Algeria TV . [1 1 PARIS 1

]
1996-08-22 [2 2 Algerian 2] security forces have shot dead
three [3 3 Moslem 3] guerrillas suspected of killing a
leading [4 4 French 4] bishop in western [5 5 Algeria 5] ,
the [6 6 Algerian 6] state-run television said on Thursday .
Security forces also arrested four other men sought for
giving support to the slain [7 7 Moslem 7] rebels , the
television said . The television , which did not say when the
security forces killed the rebels , said the four arrested
men confessed details of the assassination of the [8 8 French
8] [9 9 Roman Catholic 9] Bishop Pierre Claverie . The 58-
year-old Claverie was killed in August 1 in a bomb blast at
his residence in the western [10 10 Algerian 10] city of [11 1
1 Oran 11] , hours after he met visiting [12 12 French 12]
Foreign Minister [13 13 Herve de Charette 13] in [14 14 Algiers
14] . An estimated 50,000 [1 Algerians 1] and more than
110 foreigners have been killed in [15 15 Algeria 15] 's
violence pitting [16 16 Moslem 16] rebels against the [17 17 Algerian
17] government forces since early 1992 , when the
authorities cancelled a general election in which radical [2
Islamists 2] took a commanding lead .

1: 22989 Paris

2: 358 Algeria

3: 6037917 Islam

4: 5843419 France

5: 358 Algeria

6: 358 Algeria

7: 6037917 Islam

8: 5843419 France

9: 606848 Catholic_Church

10: 358 Algeria

11: 145335 Oran

12: 5843419 France

13: 595568 Hervé_de_Charette

14: 1644 Algiers

15: 358 Algeria

16: 6037917 Islam

17: 358 Algeria

1: 22989 Paris

[pem_pos=0, final_scores=0.5157785,
final_local_scores=0.5273089, pem=-0.115301244,
similarity=20.452688]

2: 358 Algeria

[pem_pos=0, final_scores=0.78218955,
final_local_scores=0.75871396, pem=-0.20917186,
similarity=27.86876]

3: 6037917 Islam

[pem_pos=1, final_scores=0.38411778,
final_local_scores=0.40672404, pem=-1.2150909,
similarity=22.835022]

4: 5843419 France

[pem_pos=0, final_scores=0.7095972,
final_local_scores=0.6940044, pem=-0.80223274,

```
similarity=29.162855]
5: 358 Algeria
    [pem_pos=0, final_scores=0.5719302,
final_local_scores=0.5290376, pem=-0.33694586,
similarity=21.709614]
6: 358 Algeria
    [pem_pos=0, final_scores=0.60409695,
final_local_scores=0.5641749, pem=-0.20917186,
similarity=22.063326]
7: 6037917 Islam
    [pem_pos=1, final_scores=0.41994447,
final_local_scores=0.4458593, pem=-1.2150909,
similarity=24.002897]
8: 5843419 France
    [pem_pos=0, final_scores=0.71096534,
final_local_scores=0.6954988, pem=-0.80223274,
similarity=29.20745]
9: 606848 Catholic_Church
    [pem_pos=0, final_scores=0.38170487,
final_local_scores=0.42191565, pem=0.0, similarity=16.680517]
10: 358 Algeria
    [pem_pos=0, final_scores=0.6043632,
final_local_scores=0.5644657, pem=-0.20917186,
similarity=22.072004]
11: 145335 Oran
    [pem_pos=0, final_scores=0.40674788,
final_local_scores=0.39983863, pem=-0.12748516,
similarity=16.71498]
12: 5843419 France
    [pem_pos=0, final_scores=0.68245274,
final_local_scores=0.6643532, pem=-0.80223274,
similarity=28.278004]
13: 595568 Hervé_de_Charette
    [pem_pos=0, final_scores=0.55300486,
final_local_scores=0.7367134, pem=0.0, similarity=26.07471]
14: 1644 Algiers
    [pem_pos=0, final_scores=0.3546204,
final_local_scores=0.3179784, pem=-0.10460702,
similarity=14.147692]
15: 358 Algeria
    [pem_pos=0, final_scores=0.65623456,
final_local_scores=0.6211273, pem=-0.33694586,
similarity=24.457754]
16: 6037917 Islam
    [pem_pos=1, final_scores=0.40655226,
final_local_scores=0.43123037, pem=-1.2150909,
similarity=23.56634]
17: 358 Algeria
    [pem_pos=0, final_scores=0.72097903,
final_local_scores=0.6918508, pem=-0.20917186,
similarity=25.873428]
1: 358 Algeria
```

```
[pem_pos=0, final_scores=0.23487882,
final_local_scores=0.16086015, pem=0.0, similarity=8.8901005]
2: 15012 Islamism
[pem_pos=0, final_scores=0.034373507,
final_local_scores=0.04380171, pem=0.0, similarity=5.3968434]
```

Sample: b'105_German'

```
[1 1 German 1] flown cargo January-July rise 3.8 percent . [2
2 FRANKFURT 2] 1996-08-22 The following table shows total
flown air cargo volumes in tonnes handled at international [3
3 German 3] airports January-July 1996 . The figures exclude
trucked airfreight according to the [4 4 German 4] airports
association ADV . [5 5 Berlin 5] ( total ) 17,844 up 5.9 pct
- [6 6 Tegel 6] 10,896 up 3.1 - [7 7 Tempelhof 7] 202 down
60.0 - Schoenefeld 6,746 up 16.8 [8 8 Bremen 8] 1,453 up 13.1
1 Dresden 792 up 11.4 Duesseldorf 31,347 down 4.4 [9 9 Frank
furt 9] 768,269 up 1.5 [10 10 Hamburg 10] 21,240 down 3.5
Hannover 6,030 up 15.3 Koeln ( Cologne ) 182,887 up 11.8 [11 1
1 Leipzig 11] / Halle 1,806 up 45.6 [12 12 Munich 12] 44,525
up 11.8 Muenster / Osnabruock 382 up 28.2 [13 13 Nuremberg 13]
25,929 up 17.8 Saarbruecken 626 up 28.3 [14 14 Stuttgart 14]
10,655 up 11.7 TOTAL 1,113,785 up 3.8 - Air Cargo Newsroom
Tel+44 161 542 7706 Fax+44 171 542 5017
```

```
1: 11867 Germany
2: 10992 Frankfurt
3: 11867 Germany
4: 11867 Germany
5: 3354 Berlin
6: 271834 Berlin_Tegel_Airport
7: 271851 Berlin_Tempelhof_Airport
8: 500570 Bremen_Airport
9: 248285 Frankfurt_Airport
10: 900545 Hamburg_Airport
11: 5454926 Leipzig/Halle_Airport
12: 266998 Munich_Airport
13: 1789618 Nuremberg_Airport
14: 270015 Stuttgart_Airport
```

```
1: 917790 Dresden_Airport
[winner: 420993 SMS_Dresden_(1907), pem_pos=7,
final_scores=-0.35568833, final_local_scores=-0.31370744,
pem=-3.542201, similarity=13.9910965|
RECALL MISS: 917790 Dresden_Airport]
```

```
1: 11867 Germany
[pem_pos=1, final_scores=0.6629229,
final_local_scores=0.6718889, pem=-1.0465823,
similarity=29.831696]
2: 10992 Frankfurt
[pem_pos=0, final_scores=0.5369149,
final_local_scores=0.51786685, pem=-0.07062824,
similarity=19.92798]
3: 11867 Germany
[pem_pos=1, final_scores=0.70232064,
```

```
final_local_scores=0.7149249, pem=-1.0465823,
similarity=31.115974]
4: 11867 Germany
    [pem_pos=1, final_scores=0.6594458,
final_local_scores=0.66809064, pem=-1.0465823,
similarity=29.718348]
5: 3354 Berlin
    [pem_pos=0, final_scores=0.3337155,
final_local_scores=0.31515798, pem=-0.15483057,
similarity=14.336648]
6: 271834 Berlin_Tegel_Airport
    [pem_pos=1, final_scores=0.34059054,
final_local_scores=0.31303847, pem=-1.0794699,
similarity=19.301731]
7: 271851 Berlin_Tempelhof_Airport
    [pem_pos=1, final_scores=0.38956386,
final_local_scores=0.37617427, pem=-0.23921861,
similarity=16.616411]
8: 500570 Bremen_Airport
    [pem_pos=6, final_scores=0.46587127,
final_local_scores=0.4970286, pem=-3.2606614,
similarity=36.654022]
9: 248285 Frankfurt_Airport
    [pem_pos=1, final_scores=0.38968235,
final_local_scores=0.36501545, pem=-2.9948287,
similarity=31.268845]
10: 900545 Hamburg_Airport
    [pem_pos=17, final_scores=0.36520505,
final_local_scores=0.35824454, pem=-4.059266,
similarity=36.855366]
11: 5454926 Leipzig/Halle_Airport
    [pem_pos=10, final_scores=0.5180995,
final_local_scores=0.53763986, pem=-2.9177907,
similarity=36.001358]
12: 266998 Munich_Airport
    [pem_pos=10, final_scores=0.4114684,
final_local_scores=0.3960445, pem=-3.5990505,
similarity=35.480667]
13: 1789618 Nuremberg_Airport
    [pem_pos=10, final_scores=0.36714,
final_local_scores=0.38215244, pem=-2.9599967,
similarity=31.590828]
14: 270015 Stuttgart_Airport
    [pem_pos=13, final_scores=0.34545618,
final_local_scores=0.3457511, pem=-3.1169467,
similarity=31.358055]
```

Sample: b'106_Paribas'
[1 1 Paribas 1] repeats buy on [2 2 Aegon 2] after results .
[3 3 AMSTERDAM 3] 1996-08-22 Summary of Aug 22 research .
Company-----Price---Broker----- [4 4 Aegon
4] 83.40 [5 5 Paribas 5] COMMENT : " Not only did [6 6 Aegon 6

] surprise with earnings of 711 million guilders , which were above the top of the expected range , it also forecast a similar performance in the second half . " Reiterates previous " buy " recommendation after results . Estimates (Dfl) : EPS P / E Dividend 1996 5.83 13.8 2.75 1997 6.59 12.2 3.10 -- [7 7 Amsterdam 7] newsroom , +31 20 504 5000 (Fax +31 20 504 5040)
1: 564293 BNP_Paribas
2: 2186512 Aegon
3: 844 Amsterdam
4: 2186512 Aegon
5: 564293 BNP_Paribas
6: 2186512 Aegon
7: 844 Amsterdam
1: 564293 BNP_Paribas
[pem_pos=0, final_scores=0.41240662,
final_local_scores=0.416106, pem=0.0, similarity=16.507145]
2: 2186512 Aegon
[pem_pos=0, final_scores=0.5221312,
final_local_scores=0.5042288, pem=0.0, similarity=19.136906]
3: 844 Amsterdam
[pem_pos=0, final_scores=0.61991674,
final_local_scores=0.6738308, pem=-0.0642803,
similarity=24.547735]
4: 2186512 Aegon
[pem_pos=0, final_scores=0.40320086,
final_local_scores=0.37431544, pem=0.0, similarity=15.260032]
5: 564293 BNP_Paribas
[pem_pos=0, final_scores=0.40649945,
final_local_scores=0.40965325, pem=0.0, similarity=16.314583]
6: 2186512 Aegon
[pem_pos=0, final_scores=0.5051394,
final_local_scores=0.48566782, pem=0.0, similarity=18.58301]
7: 844 Amsterdam
[pem_pos=0, final_scores=0.59099334,
final_local_scores=0.6422365, pem=-0.0642803,
similarity=23.604897]

Sample: b'107_Clinton'

[1 1 Clinton 1] 's [2 2 Ballybunion 2] fans invited to [3 3 C hicago 3] . [4 4 DUBLIN 4] 1996-08-22 [5 5 U.S. 5] President [6 6 Bill Clinton 6] had to drop the resort of [7 7 Ballybunion 7] from a whirlwind [8 8 Irish 8] tour last year . So [9 9 Ba llybunion 9] is going to [10 10 America 10] instead . Two residents of the Atlantic resort , where [11 11 Clinton 11] was to have played golf with the [12 12 Irish 12] Foreign Minister [13 13 Dick Spring 13] , have been invited to the [14 14 Democratic 14] party convention in [15 15 Chicago 15] on August 26-29 . They have been asked to bring with them the placards they waved when [16 16 Clinton 16] addressed [17 17 I reland 17] at a packed ceremony in [18 18 Dublin 18] city centre on December 1 , last year . They read : " [19 19 Bally

bunion 19] backs [20 20 Clinton 20] . " " The [21 21 Democrat
ic 21] party have requested we bring our placards with us .
We will be guests of the [22 22 Kennedys 22] , " said Frank
Quilter , one of the two who have been invited to [23 23 Chic
ago 23] . [24 24 Clinton 24] made a triumphant [25 25 Irish 25
] tour to back a [26 26 Northern Ireland 26] peace process
but was forced to drop [27 27 Ballybunion 27] from a packed
schedule at the last minute .

1: 3356 Bill_Clinton
2: 991564 Ballybunion
3: 6886 Chicago
4: 8504 Dublin
5: 3434750 United_States
6: 3356 Bill_Clinton
7: 991564 Ballybunion
8: 14560 Republic_of_Ireland
9: 991564 Ballybunion
10: 3434750 United_States
11: 3356 Bill_Clinton
12: 14560 Republic_of_Ireland
13: 264116 Dick_Spring
14: 5043544 Democratic_Party_(United_States)
15: 6886 Chicago
16: 3356 Bill_Clinton
17: 14560 Republic_of_Ireland
18: 8504 Dublin
19: 991564 Ballybunion
20: 3356 Bill_Clinton
21: 5043544 Democratic_Party_(United_States)
22: 7957 Kennedy_family
23: 6886 Chicago
24: 3356 Bill_Clinton
25: 14560 Republic_of_Ireland
26: 21265 Northern_Ireland
27: 991564 Ballybunion
1: 3356 Bill_Clinton
 [pem_pos=0, final_scores=0.7644114,
final_local_scores=0.768249, pem=0.0, similarity=27.015795]
2: 991564 Ballybunion
 [pem_pos=0, final_scores=0.46188712,
final_local_scores=0.5233782, pem=0.0, similarity=19.708363]
3: 6886 Chicago
 [pem_pos=0, final_scores=0.67639375,
final_local_scores=0.68604857, pem=-0.20658566,
similarity=25.686214]
4: 8504 Dublin
 [pem_pos=0, final_scores=0.5127747,
final_local_scores=0.50693905, pem=-0.23150389,
similarity=20.47674]
5: 3434750 United_States
 [pem_pos=0, final_scores=0.50623184,
final_local_scores=0.49950093, pem=-0.14075686,

```
similarity=19.761276]
6: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.6037306,
final_local_scores=0.58822405, pem=0.0, similarity=21.643492]
7: 991564 Ballybunion
    [pem_pos=0, final_scores=0.50071555,
final_local_scores=0.5657924, pem=0.0, similarity=20.974087]
8: 14560 Republic_of_Ireland
    [pem_pos=3, final_scores=0.5444308,
final_local_scores=0.55456597, pem=-2.2576945,
similarity=32.91676]
9: 991564 Ballybunion
    [pem_pos=0, final_scores=0.43034214,
final_local_scores=0.4889201, pem=0.0, similarity=18.680063]
10: 3434750 United_States
    [pem_pos=0, final_scores=0.4769953,
final_local_scores=0.4675644, pem=-0.78483206,
similarity=22.310806]
11: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.8418445,
final_local_scores=0.85283285, pem=0.0, similarity=29.539946]
12: 14560 Republic_of_Ireland
    [pem_pos=3, final_scores=0.5236072,
final_local_scores=0.5318194, pem=-2.2576945,
similarity=32.237957]
13: 264116 Dick_Spring
    [pem_pos=0, final_scores=0.55927354,
final_local_scores=0.6281398, pem=0.0, similarity=22.834658]
14: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.44378752,
final_local_scores=0.45163858, pem=-0.076093495,
similarity=17.981318]
15: 6886 Chicago
    [pem_pos=0, final_scores=0.60357064,
final_local_scores=0.6065003, pem=-0.20658566,
similarity=23.312338]
16: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.80192417,
final_local_scores=0.80922604, pem=0.0, similarity=28.238632]
17: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.42562085,
final_local_scores=0.42478412, pem=-1.5040284,
similarity=24.94526]
18: 8504 Dublin
    [pem_pos=0, final_scores=0.41819334,
final_local_scores=0.40362328, pem=-0.23150389,
similarity=17.39359]
19: 991564 Ballybunion
    [pem_pos=0, final_scores=0.39865345,
final_local_scores=0.454305, pem=0.0, similarity=17.64708]
20: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.75812525,
```

```
final_local_scores=0.76138234, pem=0.0, similarity=26.81088]
21: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.37777382,
final_local_scores=0.37952864, pem=-0.076093495,
similarity=15.829414]
22: 7957 Kennedy_family
    [pem_pos=0, final_scores=0.3189577,
final_local_scores=0.32793444, pem=0.0, similarity=13.875931]
23: 6886 Chicago
    [pem_pos=0, final_scores=0.6002696,
final_local_scores=0.6028945, pem=-0.20658566,
similarity=23.204733]
24: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.83423907,
final_local_scores=0.84452504, pem=0.0, similarity=29.292025]
25: 14560 Republic_of_Ireland
    [pem_pos=3, final_scores=0.5622079,
final_local_scores=0.5739847, pem=-2.2576945,
similarity=33.496254]
26: 21265 Northern_Ireland
    [pem_pos=0, final_scores=0.4031797,
final_local_scores=0.39388442, pem=-0.25044483,
similarity=17.205967]
27: 991564 Ballybunion
    [pem_pos=0, final_scores=0.42800605,
final_local_scores=0.4863683, pem=0.0, similarity=18.603912]
```

Sample: b'108_Bonn'

Bonn says 1 Moscow has promised to observe ceasefire . [1 1 B
ONN 1] 1996-08-22 [2 2 Germany 2] said on Thursday it had
received assurances from the [3 3 Russian 3] government that
its forces would observe the latest ceasefire in [4 4 Chechnya
4] . [5 5 Foreign Ministry 5] spokesman Martin Erdmann said
top [6 6 Bonn 6] diplomat [7 7 Wolfgang Ischinger 7] had been
assured by senior [8 8 Russian 8] officials that the
ultimatum to storm and take the [9 9 Chechen 9] capital of [10
10 Grozny 10] was not valid . " The [11 11 Russian 11] side
confirmed that the ceasefire is in place and they will keep
to it , " Erdmann told [12 12 Reuters 12] after speaking by
telephone to [13 13 Ischinger 13] , who had met the officials
on a two-day visit to [14 14 Moscow 14] . He returned to [15 1
5 Bonn 15] on Thursday . [16 16 Ischinger 16] is the
political director of [17 17 Bonn 17] 's foreign ministry . [1
8 18 Ischinger 18] said he met three [19 19 Russian 19]
deputy foreign ministers and a vice defence minister , who
confirmed [20 20 Russian 20] Foreign Minister [21 21 Yevgeny P
rimakov 21] 's pledge that 2 Moscow would seek a political
solution under the aegis of the Organisation for Security and
Cooperation in Europe (OSCE) . " The ultimatum (to storm [2
2 22 Grozny 22]) is no longer an issue , " he said quoting [2
3 23 Ischinger 23] , who had been sent to [24 24 Moscow 24]
by [25 25 German 25] Foreign Minister [26 26 Klaus Kinkel 26]

as his personal envoy to urge an end to 3 Moscow 's military campaign in the breakaway region . [27 27 Ischinger 27] said the threat of a major assault to take [28 28 Grozny 28] had been the unauthorised initiative of the commanding general and not 4 Moscow 's intention . The officials had been positive about [29 29 Kinkel 29] 's request on Wednesday that President [30 30 Boris Yeltsin 30] 's security chief [31 31 Alexander Lebed 31] should , on his return to [32 32 Moscow 32] , meet Tim Goldiman , the OSCE representative responsible for [33 33 Chechnya 33] , he said .

1: 3295 Bonn
2: 11867 Germany
3: 25391 Russia
4: 6095 Chechnya
5: 2253457 Foreign_Office_(Germany)
6: 3295 Bonn
7: 72676 Wolfgang_Ischinger
8: 25391 Russia
9: 6095 Chechnya
10: 140790 Grozny
11: 25391 Russia
12: 18998750 Reuters
13: 72676 Wolfgang_Ischinger
14: 19004 Moscow
15: 3295 Bonn
16: 72676 Wolfgang_Ischinger
17: 3295 Bonn
18: 72676 Wolfgang_Ischinger
19: 25391 Russia
20: 25391 Russia
21: 520754 Yevgeny_Primakov
22: 140790 Grozny
23: 72676 Wolfgang_Ischinger
24: 19004 Moscow
25: 11867 Germany
26: 666485 Klaus_Kinkel
27: 72676 Wolfgang_Ischinger
28: 140790 Grozny
29: 666485 Klaus_Kinkel
30: 23278520 Boris_Yeltsin
31: 431579 Alexander_Lebed
32: 19004 Moscow
33: 6095 Chechnya
1: 25391 Russia
[winner: 19004 Moscow, pem_pos=0,
final_scores=-0.25710064, final_local_scores=-0.3590623,
pem=-0.106582455, similarity=-6.0458107|
RECALL MISS: 25391 Russia]
2: 25391 Russia
[winner: 19004 Moscow, pem_pos=0,
final_scores=-0.39257556, final_local_scores=-0.5070481,
pem=-0.106582455, similarity=-10.462001|

RECALL MISS: 25391 Russia]
3: 25391 Russia
[winner: 19004 Moscow, pem_pos=0,
final_scores=-0.10100472, final_local_scores=-0.18855125,
pem=-0.106582455, similarity=-0.9574206]
RECALL MISS: 25391 Russia]
4: 25391 Russia
[winner: 19004 Moscow, pem_pos=0,
final_scores=-0.18795894, final_local_scores=-0.2835355,
pem=-0.106582455, similarity=-3.7919405]
RECALL MISS: 25391 Russia]
1: 3295 Bonn
[pem_pos=0, final_scores=0.45450145,
final_local_scores=0.43400246, pem=0.0, similarity=17.041212]
2: 11867 Germany
[pem_pos=0, final_scores=0.67968494,
final_local_scores=0.6930322, pem=-0.15529114,
similarity=25.615675]
3: 25391 Russia
[pem_pos=1, final_scores=0.87731165,
final_local_scores=0.9113844, pem=-1.0515022,
similarity=37.005474]
4: 6095 Chechnya
[pem_pos=0, final_scores=0.5248166,
final_local_scores=0.52982354, pem=0.0, similarity=19.900703]
5: 2253457 Foreign_Office_(Germany)
[pem_pos=11, final_scores=0.060900018,
final_local_scores=0.047533393, pem=-2.984704,
similarity=21.739489]
6: 3295 Bonn
[pem_pos=0, final_scores=0.41581726,
final_local_scores=0.39174587, pem=0.0, similarity=15.78019]
7: 72676 Wolfgang_Ischinger
[pem_pos=0, final_scores=0.690219,
final_local_scores=0.7574349, pem=0.0, similarity=26.69308]
8: 25391 Russia
[pem_pos=1, final_scores=0.8785402,
final_local_scores=0.91272646, pem=-1.0515022,
similarity=37.045525]
9: 6095 Chechnya
[pem_pos=2, final_scores=0.50486237,
final_local_scores=0.50802654, pem=-0.8617233,
similarity=23.936424]
10: 140790 Grozny
[pem_pos=0, final_scores=0.5226105,
final_local_scores=0.5357808, pem=0.0, similarity=20.07848]
11: 25391 Russia
[pem_pos=1, final_scores=0.82531995,
final_local_scores=0.85459137, pem=-1.0515022,
similarity=35.310658]
12: 18998750 Reuters
[pem_pos=0, final_scores=0.35637945,

```
final_local_scores=0.3912061, pem=0.0, similarity=15.764082]
13: 72676 Wolfgang_Ischinger
    [pem_pos=0, final_scores=0.5418422,
final_local_scores=0.59535587, pem=0.0, similarity=21.85632]
14: 19004 Moscow
    [pem_pos=0, final_scores=0.5138511,
final_local_scores=0.48695368, pem=-0.106582455,
similarity=19.200994]
15: 3295 Bonn
    [pem_pos=0, final_scores=0.46136123,
final_local_scores=0.44149578, pem=0.0, similarity=17.264828]
16: 72676 Wolfgang_Ischinger
    [pem_pos=0, final_scores=0.5188002,
final_local_scores=0.57018596, pem=0.0, similarity=21.1052]
17: 3295 Bonn
    [pem_pos=0, final_scores=0.45885038,
final_local_scores=0.438753, pem=0.0, similarity=17.182978]
18: 72676 Wolfgang_Ischinger
    [pem_pos=0, final_scores=0.43771893,
final_local_scores=0.48161697, pem=0.0, similarity=18.462124]
19: 25391 Russia
    [pem_pos=1, final_scores=0.8428547,
final_local_scores=0.8737454, pem=-1.0515022,
similarity=35.88225]
20: 25391 Russia
    [pem_pos=1, final_scores=0.9278776,
final_local_scores=0.96662, pem=-1.0515022,
similarity=38.653816]
21: 520754 Yevgeny_Primakov
    [pem_pos=0, final_scores=0.72896993,
final_local_scores=0.80252576, pem=0.0, similarity=28.038683]
22: 140790 Grozny
    [pem_pos=0, final_scores=0.5811233,
final_local_scores=0.5996971, pem=0.0, similarity=21.98587]
23: 72676 Wolfgang_Ischinger
    [pem_pos=0, final_scores=0.5497682,
final_local_scores=0.6040138, pem=0.0, similarity=22.114689]
24: 19004 Moscow
    [pem_pos=0, final_scores=0.4712134,
final_local_scores=0.4403785, pem=-0.106582455,
similarity=17.811098]
25: 11867 Germany
    [pem_pos=1, final_scores=0.81488246,
final_local_scores=0.8407149, pem=-1.0465823,
similarity=34.8698]
26: 666485 Klaus_Kinkel
    [pem_pos=0, final_scores=0.72253126,
final_local_scores=0.81621915, pem=0.0, similarity=28.44732]
27: 72676 Wolfgang_Ischinger
    [pem_pos=0, final_scores=0.5045224,
final_local_scores=0.55458957, pem=0.0, similarity=20.639772]
28: 140790 Grozny
```

```
[pem_pos=0, final_scores=0.5447197,
final_local_scores=0.5599317, pem=0.0, similarity=20.799192]
29: 666485 Klaus_Kinkel
    [pem_pos=0, final_scores=0.50816745,
final_local_scores=0.5820592, pem=0.0, similarity=21.45952]
30: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.58283216,
final_local_scores=0.6019327, pem=0.0, similarity=22.052586]
31: 431579 Alexander_Lebed
    [pem_pos=0, final_scores=0.66375583,
final_local_scores=0.7401797, pem=0.0, similarity=26.178152]
32: 19004 Moscow
    [pem_pos=0, final_scores=0.56812495,
final_local_scores=0.5462397, pem=-0.106582455,
similarity=20.970207]
33: 6095 Chechnya
    [pem_pos=0, final_scores=0.5167122,
final_local_scores=0.5209707, pem=0.0, similarity=19.636518]
```

Sample: b'109_India'

[1 1 India 1] says sees no arms race with [2 2 China 2] , [3 3 Pakistan 3] . [4 4 NEW DELHI 4] 1996-08-22 [5 5 India 5] said on Thursday that its opposition to a global nuclear test ban treaty did not mean New Delhi intended to enter into an arms race with neighbouring [6 6 Pakistan 6] and [7 7 China 7] . Foreign Minister [8 8 I.K. Gujral 8] was asked at a news conference if [9 9 India 9] 's decision to block adoption of the accord in [10 10 Geneva 10] would lead to an arms race with [11 11 Pakistan 11] and [12 12 China 12] . " I do n't see that possibility because [13 13 India 13] is not entering into any arms race , " he said . " Our not signing a new treaty does not mean we are going in for any new kind of weapons , particularly nuclear . " [14 14 China 14] , along with [15 15 Britain 15] , [16 16 France 16] , [17 17 Russia 17] and the [18 18 United States 18] , is a declared nuclear power . [19 19 India 19] carried out a nuclear test in 1974 but says it has not built the bomb . Experts believe both [20 20 India 20] and [21 21 Pakistan 21] could quickly assemble nuclear weapons . [22 22 Gujral 22] said he did not expect [23 23 India 23] 's veto of the [24 24 Comprehensive Test Ban Treaty 24] ([25 25 CTBT 25]) to damage bilateral ties with other nations . " I do not visualise its straining our bilateral relations with any country . The text has already been blocked , " he said . [26 26 Gujral 26] said [27 27 India 27] would re-examine its position if the treaty , particularly a clause providing for its entry into force , was modified . Asked what [28 28 India 28] would do if the pact were forwarded to the United Nations General Assembly , [29 29 Gujral 29] said : " That bridge I will cross when I come to it . " In a written statement released at the news conference , [30 30 Gujral 30] reiterated [31 31 India 31] 's objections to the treaty , under negotiation at the

Conference on Disarmament in [32 32 Geneva 32] . " It is a sad fact that the nuclear weapon states show no interest in giving up their nuclear hegemony , " the statement said . [33 33 Gujral 33] said [34 34 India 34] had national security concerns that made it impossible for New Delhi to sign the [35 35 CTBT 35] . " Our security concerns oblige us to maintain our nuclear option , " he said , adding that [36 36 India 36] had exercised restraint in not carrying out any nuclear tests since the country 's lone test blast in 1974 . He said : " We cannot accept constraints on our option as long as nuclear weapon states continue to rely on their nuclear arsenals for their security " .

1: 14533 India
2: 5405 China
3: 23235 Pakistan
4: 51585 New_Delhi
5: 14533 India
6: 23235 Pakistan
7: 5405 China
8: 373750 I._K._Gujral
9: 14533 India
10: 12521 Geneva
11: 23235 Pakistan
12: 5405 China
13: 14533 India
14: 5405 China
15: 31717 United_Kingdom
16: 5843419 France
17: 25391 Russia
18: 3434750 United_States
19: 14533 India
20: 14533 India
21: 23235 Pakistan
22: 373750 I._K._Gujral
23: 14533 India
24: 7851 Comprehensive_Nuclear-Test-Ban_Treaty
25: 7851 Comprehensive_Nuclear-Test-Ban_Treaty
26: 373750 I._K._Gujral
27: 14533 India
28: 14533 India
29: 373750 I._K._Gujral
30: 373750 I._K._Gujral
31: 14533 India
32: 12521 Geneva
33: 373750 I._K._Gujral
34: 14533 India
35: 7851 Comprehensive_Nuclear-Test-Ban_Treaty
36: 14533 India
1: 14533 India
[pem_pos=0, final_scores=0.705468,
final_local_scores=0.67732507, pem=-0.13622136,
similarity=25.043238]

```
2: 5405 China
    [pem_pos=0, final_scores=0.48604053,
final_local_scores=0.48162317, pem=-0.15669827,
similarity=19.314459]
3: 23235 Pakistan
    [pem_pos=0, final_scores=0.5748634,
final_local_scores=0.5504075, pem=-0.21944217,
similarity=21.708332]
4: 51585 New_Delhi
    [pem_pos=0, final_scores=0.40859634,
final_local_scores=0.39402974, pem=0.0, similarity=15.848347]
5: 14533 India
    [pem_pos=0, final_scores=0.73055357,
final_local_scores=0.70472735, pem=-0.13622136,
similarity=25.860975]
6: 23235 Pakistan
    [pem_pos=0, final_scores=0.5455428,
final_local_scores=0.5183792, pem=-0.21944217,
similarity=20.752542]
7: 5405 China
    [pem_pos=0, final_scores=0.5424948,
final_local_scores=0.54329085, pem=-0.15669827,
similarity=21.154745]
8: 373750 I._K._Gujral
    [pem_pos=0, final_scores=0.7938631,
final_local_scores=0.8453844, pem=0.0, similarity=29.31767]
9: 14533 India
    [pem_pos=0, final_scores=0.66453236,
final_local_scores=0.6326091, pem=-0.13622136,
similarity=23.708824]
10: 12521 Geneva
    [pem_pos=0, final_scores=0.46738666,
final_local_scores=0.5047038, pem=-0.24772356,
similarity=20.498241]
11: 23235 Pakistan
    [pem_pos=0, final_scores=0.6072721,
final_local_scores=0.58580905, pem=-0.21944217,
similarity=22.764784]
12: 5405 China
    [pem_pos=0, final_scores=0.52982134,
final_local_scores=0.529447, pem=-0.15669827,
similarity=20.741617]
13: 14533 India
    [pem_pos=0, final_scores=0.5834662,
final_local_scores=0.54405665, pem=-0.13622136,
similarity=21.066242]
14: 5405 China
    [pem_pos=0, final_scores=0.44747084,
final_local_scores=0.4394917, pem=-0.15669827,
similarity=18.05717]
15: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.55380064,
```

```
final_local_scores=0.5600565, pem=-0.61114657,
similarity=24.126425]
16: 5843419 France
    [pem_pos=0, final_scores=0.63767713,
final_local_scores=0.6831837, pem=-0.16240834,
similarity=25.36048]
17: 25391 Russia
    [pem_pos=0, final_scores=0.69881326,
final_local_scores=0.74480516, pem=-0.20299491,
similarity=27.420101]
18: 3434750 United_States
    [pem_pos=0, final_scores=0.51309305,
final_local_scores=0.5189209, pem=-0.073457584,
similarity=19.974821]
19: 14533 India
    [pem_pos=0, final_scores=0.70102364,
final_local_scores=0.67247033, pem=-0.13622136,
similarity=24.898361]
20: 14533 India
    [pem_pos=0, final_scores=0.67803997,
final_local_scores=0.6473642, pem=-0.13622136,
similarity=24.149145]
21: 23235 Pakistan
    [pem_pos=0, final_scores=0.57328314,
final_local_scores=0.54868126, pem=-0.21944217,
similarity=21.656816]
22: 373750 I._K._Gujral
    [pem_pos=0, final_scores=0.68852305,
final_local_scores=0.7362305, pem=0.0, similarity=26.060299]
23: 14533 India
    [pem_pos=0, final_scores=0.72101647,
final_local_scores=0.6943095, pem=-0.13622136,
similarity=25.550085]
24: 7851 Comprehensive_Nuclear-Test-Ban_Treaty
    [pem_pos=0, final_scores=0.41694546,
final_local_scores=0.51333016, pem=0.0, similarity=19.408508]
25: 7851 Comprehensive_Nuclear-Test-Ban_Treaty
    [pem_pos=0, final_scores=0.6871711,
final_local_scores=0.80851066, pem=0.0, similarity=28.217283]
26: 373750 I._K._Gujral
    [pem_pos=0, final_scores=0.7490538,
final_local_scores=0.8023511, pem=0.0, similarity=28.03347]
27: 14533 India
    [pem_pos=0, final_scores=0.687344,
final_local_scores=0.6575274, pem=-0.13622136,
similarity=24.452435]
28: 14533 India
    [pem_pos=0, final_scores=0.6090556,
final_local_scores=0.57200927, pem=-0.13622136,
similarity=21.900404]
29: 373750 I._K._Gujral
    [pem_pos=0, final_scores=0.71612144,
```

```
final_local_scores=0.7663775, pem=0.0, similarity=26.959946]
30: 373750 I.K.Gujral
    [pem_pos=0, final_scores=0.66178715,
final_local_scores=0.7070256, pem=0.0, similarity=25.188766]
31: 14533 India
    [pem_pos=0, final_scores=0.62246007,
final_local_scores=0.58665156, pem=-0.13622136,
similarity=22.337358]
32: 12521 Geneva
    [pem_pos=0, final_scores=0.45664686,
final_local_scores=0.4929722, pem=-0.24772356,
similarity=20.148146]
33: 373750 I.K.Gujral
    [pem_pos=0, final_scores=0.74351746,
final_local_scores=0.7963035, pem=0.0, similarity=27.852999]
34: 14533 India
    [pem_pos=0, final_scores=0.75516766,
final_local_scores=0.7316144, pem=-0.13622136,
similarity=26.66334]
35: 7851 Comprehensive_Nuclear-Test-Ban_Treaty
    [pem_pos=0, final_scores=0.5884053,
final_local_scores=0.70062405, pem=0.0, similarity=24.997732]
36: 14533 India
    [pem_pos=0, final_scores=0.6598411,
final_local_scores=0.6274847, pem=-0.13622136,
similarity=23.5559]
```

Sample: b'110_Britain'

[1 1 Britain 1] says death of its citizen will sour ties . [2 2 DHAKA 2] 1996-08-22 A [3 3 British 3] minister expressed his government 's official disquiet on Thursday at the recent death of a [4 4 British 4] citizen of [5 5 Bangladeshi 5] origin at [6 6 Dhaka 6] airport . " I have told [7 7 Banglade sh 7] leaders that [8 8 British 8] goverment has attached serious importance to the resolution of the tragic death of Siraj Mia , " Under-Secretary of State for Foreign and [9 9 Co mmonwealth 9] Affairs Liam Fox Fox , told reporters . Siraj Mia died at [10 10 Dhaka 10] airport on May 9 during interogation by customs officials after arriving from [11 11 L ondon 11] . His body bore multiple injuries , and his relatives complained that he was murdered . A post-mortem report suggested he might have been tortured . But customs authorities said the passenger was drunk and died of loss of blood from a deep cut in his wrist after he hit a glass sheet . [12 12 Fox 12] , who arrived in [13 13 Bangladesh 13] on Tuesday on four-day visit , said [14 14 Britain 14] wanted Dhaka to act seriously on the case . " This is one of the reasons of my visit here ... this is an important issue in our relationship " , said [15 15 Fox 15] , who is due to leave for [16 16 Nepal 16] on Friday . [17 17 Fox 17] said the incident had strained relations between the two governments . He said the Mia 's issue had been raised in the

House of Commons . [18 18 Fox 18] said he had brought up the issue at every meeting he had had with government leaders in [19 19 Dhaka 19] . He said the [20 20 Bangladesh 20] government had assured him it was taking the matter seriously . " The [21 21 British 21] government wants a thorough investigation and a just outcome , " he said . [22 22 Fox 22] said the [23 23 British 23] government wanted an end to the alleged harassment of its nationals at [24 24 Dhaka 24] airport by customs officials . [25 25 Bangladesh 25] 's [26 26 Criminal Investigation Department 26] has charged two immigration officials in connection with Mia 's killing . Mia , a father of five children , had a restaurant business in a [27 27 London 27] suburb .

1: 31717 United_Kingdom
2: 56656 Dhaka
3: 31717 United_Kingdom
4: 31717 United_Kingdom
5: 3454 Bangladesh
6: 56656 Dhaka
7: 3454 Bangladesh
8: 31717 United_Kingdom
9: 21175158 Commonwealth_of_Nations
10: 56656 Dhaka
11: 17867 London
12: 357375 Liam_Fox
13: 3454 Bangladesh
14: 31717 United_Kingdom
15: 357375 Liam_Fox
16: 171166 Nepal
17: 357375 Liam_Fox
18: 357375 Liam_Fox
19: 56656 Dhaka
20: 3454 Bangladesh
21: 31717 United_Kingdom
22: 357375 Liam_Fox
23: 31717 United_Kingdom
24: 56656 Dhaka
25: 3454 Bangladesh
26: 631179 Criminal_Investigation_Department
27: 17867 London
1: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5840506,
final_local_scores=0.54862195, pem=-0.61114657,
similarity=23.785196]
2: 56656 Dhaka
[pem_pos=0, final_scores=0.4277075,
final_local_scores=0.40627593, pem=-0.021151643,
similarity=16.328823]
3: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5742399,
final_local_scores=0.5379053, pem=-0.48936236,
similarity=22.803108]

```
4: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.52851224,
final_local_scores=0.48795474, pem=-0.48936236,
similarity=21.312485]
5: 3454 Bangladesh
    [pem_pos=0, final_scores=0.4392432,
final_local_scores=0.40621406, pem=-0.5810555,
similarity=19.37182]
6: 56656 Dhaka
    [pem_pos=0, final_scores=0.39735103,
final_local_scores=0.37311614, pem=-0.021151643,
similarity=15.33927]
7: 3454 Bangladesh
    [pem_pos=0, final_scores=0.558181,
final_local_scores=0.5361355, pem=-0.28971618,
similarity=21.664587]
8: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5424455,
final_local_scores=0.50317466, pem=-0.48936236,
similarity=21.766678]
9: 21175158 Commonwealth_of_Nations
    [pem_pos=0, final_scores=0.45803374,
final_local_scores=0.45672816, pem=-0.69133765,
similarity=21.478994]
10: 56656 Dhaka
    [pem_pos=0, final_scores=0.4298392,
final_local_scores=0.4086045, pem=-0.021151643,
similarity=16.398314]
11: 17867 London
    [pem_pos=0, final_scores=0.59708774,
final_local_scores=0.5900882, pem=-0.09086737,
similarity=22.193272]
12: 357375 Liam_Fox
    [pem_pos=13, final_scores=0.9218169,
final_local_scores=1.0210657, pem=-4.8542123,
similarity=60.958317]
13: 3454 Bangladesh
    [pem_pos=0, final_scores=0.5185589,
final_local_scores=0.49285448, pem=-0.28971618,
similarity=20.372995]
14: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.54997844,
final_local_scores=0.51140326, pem=-0.61114657,
similarity=22.674515]
15: 357375 Liam_Fox
    [pem_pos=13, final_scores=0.75506216,
final_local_scores=0.8389114, pem=-4.8542123,
similarity=55.522465]
16: 171166 Nepal
    [pem_pos=0, final_scores=0.60332483,
final_local_scores=0.63294834, pem=-0.22887385,
similarity=24.222805]
```

```
17: 357375 Liam_Fox
    [pem_pos=13, final_scores=0.8775037,
final_local_scores=0.9726601, pem=-4.8542123,
similarity=59.51379]
18: 357375 Liam_Fox
    [pem_pos=13, final_scores=0.8342791,
final_local_scores=0.92544395, pem=-4.8542123,
similarity=58.104767]
19: 56656 Dhaka
    [pem_pos=0, final_scores=0.389574,
final_local_scores=0.36462092, pem=-0.021151643,
similarity=15.085754]
20: 3454 Bangladesh
    [pem_pos=0, final_scores=0.62921345,
final_local_scores=0.61372775, pem=-0.28971618,
similarity=23.980095]
21: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.6088192,
final_local_scores=0.5756779, pem=-0.48936236,
similarity=23.930317]
22: 357375 Liam_Fox
    [pem_pos=13, final_scores=0.9439855,
final_local_scores=1.0452814, pem=-4.8542123,
similarity=61.68096]
23: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.6346309,
final_local_scores=0.60387325, pem=-0.48936236,
similarity=24.771725]
24: 56656 Dhaka
    [pem_pos=0, final_scores=0.358751,
final_local_scores=0.33095148, pem=-0.021151643,
similarity=14.080992]
25: 3454 Bangladesh
    [pem_pos=0, final_scores=0.5567145,
final_local_scores=0.5345336, pem=-0.28971618,
similarity=21.616785]
26: 631179 Criminal_Investigation_Department
    [pem_pos=0, final_scores=0.36144358,
final_local_scores=0.43570006, pem=0.0, similarity=17.091871]
27: 17867 London
    [pem_pos=0, final_scores=0.5964158,
final_local_scores=0.5893542, pem=-0.09086737,
similarity=22.171368]
```

Sample: b'111_India'

```
[1 1 India 1] fears attempts to disrupt [2 2 Kashmir 2]
polls . [3 3 SRINAGAR 3] , [4 4 India 4] 1996-08-22 [5 5 India
5] 's Home ( interior ) Minister accused [6 6 Pakistan 6] on
on Thursday of planning to disrupt state elections in
troubled Jammu and Kashmir state . " It seems that from
across the border there is going to be a planned attempt to
disrupt the elections , " Inderjit Gupta told reporters in
```

the state capital [7 7 Srinagar 7] . The local polls next month will be the first since 1987 in the state , clamped under direct rule from [8 8 New Delhi 8] since 1990 . [9 9 India 9] has often accused [10 10 Pakistan 10] of abetting militancy in the valley , a charge Islamabad has always denied . [11 11 Gupta 11] said there might be an increase in the number of people infiltrating the [12 12 Kashmir 12] valley to create disturbance in the region . " We noticed among the people who come from across the border , there is a growing number of foreign mercenaries , " [13 13 Gupta 13] said . [14 14 India 14] and [15 15 Pakistan 15] have fought two of their three wars over the troubled region of [16 16 Kashmir 16] since independence from [17 17 Britain 17] in 1947 . Prime Minister [18 18 H.D. Deve Gowda 18] 's centre-left government hopes the elections will help restore normality and democratic rule in Jammu and Kashmir , where more than 20,000 people have died in insurgency-related violence since 1990 . Over a dozen militant groups are fighting [19 19 New Delhi 19] 's rule in the state .

1: 14533 India
2: 17337 Kashmir
3: 294148 Srinagar
4: 14533 India
5: 14533 India
6: 23235 Pakistan
7: 294148 Srinagar
8: 51585 New_Delhi
9: 14533 India
10: 23235 Pakistan
11: 7140547 Indrajit_Gupta
12: 17337 Kashmir
13: 7140547 Indrajit_Gupta
14: 14533 India
15: 23235 Pakistan
16: 17337 Kashmir
17: 31717 United_Kingdom
18: 373742 H._D._Deve_Gowda
19: 51585 New_Delhi
1: 14533 India
[pem_pos=0, final_scores=0.68307126,
final_local_scores=0.646655, pem=-0.13622136,
similarity=24.127981]
2: 17337 Kashmir
[pem_pos=0, final_scores=0.383197,
final_local_scores=0.33578733, pem=-0.072103634,
similarity=14.502388]
3: 294148 Srinagar
[pem_pos=0, final_scores=0.354771,
final_local_scores=0.33663452, pem=0.0, similarity=14.135558]
4: 14533 India
[pem_pos=0, final_scores=0.7052557,
final_local_scores=0.6708882, pem=-0.13622136,

```
similarity=24.851147]
5: 14533 India
    [pem_pos=0, final_scores=0.7258308,
final_local_scores=0.69336337, pem=-0.13622136,
similarity=25.521852]
6: 23235 Pakistan
    [pem_pos=0, final_scores=0.5789166,
final_local_scores=0.5439939, pem=-0.21944217,
similarity=21.516935]
7: 294148 Srinagar
    [pem_pos=0, final_scores=0.38746274,
final_local_scores=0.3723386, pem=0.0, similarity=15.201038]
8: 51585 New_Delhi
    [pem_pos=0, final_scores=0.35407537,
final_local_scores=0.32419574, pem=0.0, similarity=13.76436]
9: 14533 India
    [pem_pos=0, final_scores=0.6318702,
final_local_scores=0.59072566, pem=-0.13622136,
similarity=22.458939]
10: 23235 Pakistan
    [pem_pos=0, final_scores=0.56449026,
final_local_scores=0.5282354, pem=-0.21944217,
similarity=21.04667]
11: 7140547 Indrajit_Gupta
    [pem_pos=0, final_scores=0.3826958,
final_local_scores=0.45407927, pem=0.0, similarity=17.640345]
12: 17337 Kashmir
    [pem_pos=0, final_scores=0.3967381,
final_local_scores=0.35057893, pem=-0.072103634,
similarity=14.943798]
13: 7140547 Indrajit_Gupta
    [pem_pos=0, final_scores=0.46981245,
final_local_scores=0.54924095, pem=0.0, similarity=20.480158]
14: 14533 India
    [pem_pos=0, final_scores=0.6791967,
final_local_scores=0.6424227, pem=-0.13622136,
similarity=24.00168]
15: 23235 Pakistan
    [pem_pos=0, final_scores=0.47901362,
final_local_scores=0.43486506, pem=-0.21944217,
similarity=18.260313]
16: 17337 Kashmir
    [pem_pos=0, final_scores=0.3698898,
final_local_scores=0.32125127, pem=-0.072103634,
similarity=14.068603]
17: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.49596828,
final_local_scores=0.5154287, pem=-0.61114657,
similarity=22.794643]
18: 373742 H._D._Deve_Gowda
    [pem_pos=0, final_scores=0.5449824,
final_local_scores=0.61162937, pem=0.0, similarity=22.341953]
```

```
19: 51585 New_Delhi
    [pem_pos=0, final_scores=0.34790844,
final_local_scores=0.3174593, pem=0.0, similarity=13.563332]
```

Sample: b'112_Dhaka'

[1 1 Dhaka 1] stocks end up on gains by engineering , banks . [2 2 DHAKA 2] 1996-08-22 [3 3 Dhaka 3] stocks edged up on sharply higher volume as engineering and cash shares gained amid buying by both small and institutional investors , brokers said . The [4 4 Dhaka Stock Exchange 4] (DSE) all-share price index rose 8.05 points or 0.7 percent to 1,156.79 on a turnover of 146.2 million taka . Of the total 119 issues traded 71 closed higher , 44 ended lower and four remained unchanged . . [5 5 National Bank 5] rose 12.71 taka to 228.7 , Eastern Cables gained 20.37 to 677.98 and Apex Tannery lost 22.72 to 597 . Brokers said the stocks recovered early losses to edge up at close because of institutional support and short-covering ahead of Friday weekend .

```
1: 4450516 Dhaka_Stock_Exchange
2: 56656 Dhaka
3: 4450516 Dhaka_Stock_Exchange
4: 4450516 Dhaka_Stock_Exchange
5: 15156528 National_Bank_(Bangladesh)
1: 4450516 Dhaka_Stock_Exchange
    [pem_pos=8, final_scores=0.46707422,
final_local_scores=0.43739337, pem=-2.3971605,
similarity=30.178537]
2: 56656 Dhaka
    [pem_pos=0, final_scores=0.32363564,
final_local_scores=0.37840295, pem=-0.021151643,
similarity=15.497038]
3: 4450516 Dhaka_Stock_Exchange
    [pem_pos=8, final_scores=0.48740417,
final_local_scores=0.45960075, pem=-2.3971605,
similarity=30.84125]
4: 4450516 Dhaka_Stock_Exchange
    [pem_pos=0, final_scores=0.5963269,
final_local_scores=0.5785822, pem=0.0, similarity=21.35576]
5: 15156528 National_Bank_(Bangladesh)
    [pem_pos=17, final_scores=0.4823367,
final_local_scores=0.5883113, pem=-2.8332133,
similarity=37.05355]
```

Sample: b'113_India'

[1 1 India 1] [2 2 RBI 2] chief sees cut in cash reserve ratio . [3 3 NEW DELHI 3] 1996-08-22 The Reserve bank of India governor [4 4 C. Rangarajan 4] said on Thursday that he expected the cash reserve ratio (CRR) maintained by banks to be reduced over the medium term . " Over the medium term , yes , " he told [5 5 Reuters 5] after addressing industrialists in the capital . He denied having said in a

recent newspaper interview that the CRR could be raised if necessary . He said he was only trying (in that newspaper report) to explain the theoretical position on the use of the CRR by central banks to manage money supply . [6 6 Rangarajan 6] explained that the cash reserve ratio was an instrument that central banks could use to regulate money supply by reducing or increasing the ratio . But in the current context , the government stood by an earlier commitment to reduce it over a period of time , he said in response to a question . -- [7 7 New Delhi 7] newsroom , +91-11-3012024

1: 14533 India
2: 277069 Reserve_Bank_of_India
3: 51585 New_Delhi
4: 30860788 C._Rangarajan
5: 18998750 Reuters
6: 30860788 C._Rangarajan
7: 51585 New_Delhi
1: 14533 India
[pem_pos=0, final_scores=0.5730166, final_local_scores=0.61700726, pem=-0.13622136, similarity=23.243233]
2: 277069 Reserve_Bank_of_India
[pem_pos=1, final_scores=0.66927105, final_local_scores=0.71195304, pem=-1.0107183, similarity=30.832254]
3: 51585 New_Delhi
[pem_pos=0, final_scores=0.3859926, final_local_scores=0.360347, pem=0.0, similarity=14.843186]
4: 30860788 C._Rangarajan
[pem_pos=0, final_scores=0.73744124, final_local_scores=0.7653658, pem=0.0, similarity=26.929754]
5: 18998750 Reuters
[pem_pos=0, final_scores=0.30576164, final_local_scores=0.4113419, pem=0.0, similarity=16.364975]
6: 30860788 C._Rangarajan
[pem_pos=0, final_scores=0.76149434, final_local_scores=0.8262077, pem=0.0, similarity=28.7454]
7: 51585 New_Delhi
[pem_pos=0, final_scores=0.40732145, final_local_scores=0.38364553, pem=0.0, similarity=15.538461]

Sample: b'114_Two'

Two pct [1 1 India 1] current account deficit viable - [2 2 R BI 2] . [3 3 BOMBAY 3] 1996-08-22 The [4 4 Reserve Bank of India 4] Governor Chakravarty Rangarajan said on Thursday that a current account deficit of two percent of gross domestic product (GDP) was sustainable given the currrent rate of growth . " The current account deficit of around two percent of GDP is a sustainable level of deficit given the expected real growth rate and the trends in imports and exports , " [5 5 Rangarajan 5] said in an address to business leaders in [6 6

New Delhi 6] . [7 7 Rangarajan 7] said a current account deficit of two percent brought about by a 16-17 percent annual growth in exports and a 14-15 percent rise in imports along with an increase in non-debt flows could lead to a reduction in the debt-service ratio to below 20 percent over the next five years . -- [8 8 Bombay 8] newsroom +91-22-2659000

1: 14533 India
2: 277069 Reserve_Bank_of_India
3: 19189 Mumbai
4: 277069 Reserve_Bank_of_India
5: 30860788 C._Rangarajan
6: 51585 New_Delhi
7: 30860788 C._Rangarajan
8: 19189 Mumbai
1: 14533 India
[pem_pos=0, final_scores=0.43524444, final_local_scores=0.45082426, pem=-0.13622136, similarity=18.284]
2: 277069 Reserve_Bank_of_India
[pem_pos=1, final_scores=0.69873935, final_local_scores=0.71627975, pem=-1.0107183, similarity=30.961372]
3: 19189 Mumbai
[pem_pos=0, final_scores=0.51117617, final_local_scores=0.49390888, pem=-0.24257013, similarity=20.148073]
4: 277069 Reserve_Bank_of_India
[pem_pos=0, final_scores=0.5241197, final_local_scores=0.5255342, pem=0.0, similarity=19.772701]
5: 30860788 C._Rangarajan
[pem_pos=0, final_scores=0.53011876, final_local_scores=0.56871295, pem=-1.609438, similarity=29.813616]
6: 51585 New_Delhi
[pem_pos=0, final_scores=0.42832237, final_local_scores=0.40136218, pem=0.0, similarity=16.067162]
7: 30860788 C._Rangarajan
[pem_pos=0, final_scores=0.42817998, final_local_scores=0.45736033, pem=-1.609438, similarity=26.490631]
8: 19189 Mumbai
[pem_pos=0, final_scores=0.40378904, final_local_scores=0.3766048, pem=-0.24257013, similarity=16.647486]

Sample: b'115_Mother'
[1 1 Mother Teresa 1] devoted to world 's poor . [2 2 CALCUTTA 2] 1996-08-22 [3 3 Mother Teresa 3] , known as the Saint of the Gutters , won the [4 4 Nobel Peace Prize 4] in 1979 for bringing hope and dignity to millions of poor , unwanted people with her simple message : " The poor must know that we

love them . " While the world heaps honours on her and even regards her as a living saint , the nun of [5 5 Albanian 5] descent maintains she is merely doing [6 6 God 6] 's work . " It gives me great joy and fulfilment to love and care for the poor and neglected , " she said . " The poor do not need our sympathy and pity . They need our love and compassion . " The diminutive [7 7 Roman Catholic 7] missionary was on respiratory support in intensive care in an [8 8 Indian 8] nursing home on Thursday after suffering heart failure . But an attending doctor said [9 9 Mother Teresa 9] , who turns 86 next Tuesday , was conscious and in stable condition . The task [10 10 Mother Teresa 10] began alone in 1949 in the slums of densely-populated [11 11 Calcutta 11] , and grew to touch the hearts of people around the world . When in 1979 she was told she had won the [12 12 Nobel Peace Prize 12] , she said characteristically : " I am unworthy . " The world disagreed , showering more than 80 national and international honours on her including the [13 13 Bharat Ratna 13] , or [1 Jewel of India 1] , the country 's highest civilian award . Her health began to deteriorate in 1989 when she was fitted with a heart pacemaker . A year later , the [14 14 Vatican 14] announced she was stepping down as Superior of her [15 15 Missionaries of Charity 15] order . More than 100 delegates flew in from around the world to elect a successor . They could not agree , so asked her to stay on . She agreed . In 1991 , [16 16 Mother Teresa 16] was treated at a [17 17 California 17] hospital for heart disease and bacterial pneumonia . In 1993 , she fell in [18 18 Rome 18] and broke three ribs . In August the same year , while in [19 19 New Delhi 19] to receive yet another award , she developed malaria , complicated by her heart and lung problems . Last April she fractured her left collar bone . But her increasing frailty , arthritis and failing eyesight has not stopped her travels around the world to mingle with the poor and desperate . [20 20 Mother Teresa 20] was born Agnes Goinxha Bejaxhiu to [21 21 Albanian 21] parents in [22 22 Skopje 22] , in what was then [23 23 Serbia 23] , on August 27 , 1910 . She attended a government school and was already deeply religious by the time she was 12 . At the age of 18 she became a [24 24 Loretto 24] nun , hoping to work at the Order 's [25 25 Calcutta 25] mission . She was sent to Loretto Abbey in [26 26 Dublin 26] and from there to [27 27 India 27] to begin her novitiate and teach geography at a convent school in [28 28 Calcutta 28] . She said her divine call to work among the poor came in September , 1946 . " The message was quite clear , " she told one interviewer . " I was to leave the convent and help the poor while living among them . It was an order . I knew where I belonged . " The [29 29 Vatican 29] and the mother superior in [30 30 Dublin 30] approved and after intensive training as a nurse with [31 31 American 31] missionaries she opened her first [32 32 Calcutta 32] slum school in December 1949 . She took the name of Teresa , after [33 33 France 33] 's Saint

Therese of the Child [34 34 Jesus 34] . In [35 35 India 35] she was simply called [36 36 Mother 36] . [37 37 Mother Teresa 37] set up her first home for the dying in a Hindu rest house in [38 38 Calcutta 38] after she saw a penniless woman turned away by a city hospital . Named " Nirmal Hriday " (Tender Heart) , it was the first of a chain of 150 homes for dying , destitute people , admitting nearly 18,000 a year . Her [39 39 Missionaries of Charity 39] , a [40 40 Roman Catholic 40] religious order she founded in 1949 , now runs about 300 homes for unwanted children and the destitute in [41 41 India 41] and abroad . In 1994 a [42 42 British 42] television documentary called the myth around [43 43 Mother Teresa 43] a mixture of " hyperbole and credulity " . [44 44 Catholics 44] around the world rose to her defence .

- 1: 347105 Mother_Teresa
- 2: 47905 Kolkata
- 3: 347105 Mother_Teresa
- 4: 26230922 Nobel_Peace_Prize
- 5: 738 Albania
- 6: 5042765 God
- 7: 606848 Catholic_Church
- 8: 14533 India
- 9: 347105 Mother_Teresa
- 10: 347105 Mother_Teresa
- 11: 47905 Kolkata
- 12: 26230922 Nobel_Peace_Prize
- 13: 236912 Bharat_Ratna
- 14: 13393 Holy_See
- 15: 345921 Missionaries_of_Charity
- 16: 347105 Mother_Teresa
- 17: 5407 California
- 18: 25458 Rome
- 19: 51585 New_Delhi
- 20: 347105 Mother_Teresa
- 21: 2927 Albanians
- 22: 29618 Skopje
- 23: 29265 Serbia
- 24: 392006 Sisters_of_Loreto
- 25: 47905 Kolkata
- 26: 8504 Dublin
- 27: 14533 India
- 28: 47905 Kolkata
- 29: 13393 Holy_See
- 30: 8504 Dublin
- 31: 3434750 United_States
- 32: 47905 Kolkata
- 33: 5843419 France
- 34: 1095706 Jesus
- 35: 14533 India
- 36: 347105 Mother_Teresa
- 37: 347105 Mother_Teresa
- 38: 47905 Kolkata

```
39: 345921 Missionaries_of_Charity
40: 606848 Catholic_Church
41: 14533 India
42: 31717 United_Kingdom
43: 347105 Mother_Teresa
44: 606848 Catholic_Church
1: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.59520835,
final_local_scores=0.5854107, pem=0.0, similarity=21.559536]
2: 47905 Kolkata
    [pem_pos=0, final_scores=0.5006549,
final_local_scores=0.48562235, pem=-0.02273471,
similarity=18.705288]
3: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.6781952,
final_local_scores=0.6760612, pem=0.0, similarity=24.264729]
4: 26230922 Nobel_Peace_Prize
    [pem_pos=0, final_scores=0.4040256,
final_local_scores=0.43349248, pem=0.0, similarity=17.025993]
5: 738 Albania
    [pem_pos=2, final_scores=0.5401109,
final_local_scores=0.577777, pem=-1.5560378,
similarity=29.793707]
6: 5042765 God
    [pem_pos=0, final_scores=0.5061093,
final_local_scores=0.5755599, pem=-0.18657415,
similarity=22.28019]
7: 606848 Catholic_Church
    [pem_pos=0, final_scores=0.44403636,
final_local_scores=0.439174, pem=0.0, similarity=17.195541]
8: 14533 India
    [pem_pos=0, final_scores=0.46199566,
final_local_scores=0.45494658, pem=-0.618294,
similarity=21.028606]
9: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.6651035,
final_local_scores=0.6617605, pem=0.0, similarity=23.837967]
10: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.7524696,
final_local_scores=0.7571947, pem=0.0, similarity=26.685913]
11: 47905 Kolkata
    [pem_pos=0, final_scores=0.4243849,
final_local_scores=0.40230894, pem=-0.02273471,
similarity=16.21905]
12: 26230922 Nobel_Peace_Prize
    [pem_pos=0, final_scores=0.33500493,
final_local_scores=0.35809785, pem=0.0, similarity=14.776067]
13: 236912 Bharat_Ratna
    [pem_pos=0, final_scores=0.43300194,
final_local_scores=0.47003144, pem=0.0, similarity=18.116388]
14: 13393 Holy_See
    [pem_pos=0, final_scores=0.27988464,
```

```
final_local_scores=0.273997, pem=-0.6812086,
similarity=15.970848]
15: 345921 Missionaries_of_Charity
    [pem_pos=0, final_scores=0.5669961,
final_local_scores=0.613755, pem=0.0, similarity=22.405386]
16: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.6146657,
final_local_scores=0.60666496, pem=0.0, similarity=22.193806]
17: 5407 California
    [pem_pos=0, final_scores=0.5102506,
final_local_scores=0.5635001, pem=-0.08662026,
similarity=21.376736]
18: 25458 Rome
    [pem_pos=0, final_scores=0.47234428,
final_local_scores=0.47662115, pem=-0.2517334,
similarity=19.682003]
19: 51585 New_Delhi
    [pem_pos=0, final_scores=0.37195492,
final_local_scores=0.3609542, pem=0.0, similarity=14.861306]
20: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.68171865,
final_local_scores=0.67991006, pem=0.0, similarity=24.379585]
21: 2927 Albanians
    [pem_pos=1, final_scores=0.15053545,
final_local_scores=0.15946433, pem=-1.2174087,
similarity=15.468904]
22: 29618 Skopje
    [pem_pos=0, final_scores=0.38725454,
final_local_scores=0.43103075, pem=0.0, similarity=16.95253]
23: 29265 Serbia
    [pem_pos=0, final_scores=0.4358163,
final_local_scores=0.49019855, pem=-0.34077588,
similarity=20.57141]
24: 392006 Sisters_of_Loreto
    [pem_pos=7, final_scores=0.3875199,
final_local_scores=0.42797464, pem=-2.2907085,
similarity=29.31856]
25: 47905 Kolkata
    [pem_pos=0, final_scores=0.42366695,
final_local_scores=0.40152472, pem=-0.02273471,
similarity=16.195646]
26: 8504 Dublin
    [pem_pos=0, final_scores=0.45198268,
final_local_scores=0.4544074, pem=-0.23150389,
similarity=18.90909]
27: 14533 India
    [pem_pos=0, final_scores=0.5736162,
final_local_scores=0.5768751, pem=-0.13622136,
similarity=22.045609]
28: 47905 Kolkata
    [pem_pos=0, final_scores=0.50603133,
final_local_scores=0.4914953, pem=-0.02273471,
```

```
similarity=18.880548]
29: 13393 Holy_See
    [pem_pos=0, final_scores=0.3045519,
final_local_scores=0.30094224, pem=-0.6812086,
similarity=16.774948]
30: 8504 Dublin
    [pem_pos=0, final_scores=0.4748841,
final_local_scores=0.47942376, pem=-0.23150389,
similarity=19.655628]
31: 3434750 United_States
    [pem_pos=0, final_scores=0.47425777,
final_local_scores=0.47849637, pem=-0.12994486,
similarity=19.07566]
32: 47905 Kolkata
    [pem_pos=0, final_scores=0.45409364,
final_local_scores=0.4347613, pem=-0.02273471,
similarity=17.187492]
33: 5843419 France
    [pem_pos=0, final_scores=0.6747099,
final_local_scores=0.7173767, pem=-0.16240834,
similarity=26.380865]
34: 1095706 Jesus
    [pem_pos=0, final_scores=0.52591187,
final_local_scores=0.5711261, pem=-0.0733903,
similarity=21.532362]
35: 14533 India
    [pem_pos=0, final_scores=0.58486944,
final_local_scores=0.58916754, pem=-0.13622136,
similarity=22.412441]
36: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.56264216,
final_local_scores=0.5498372, pem=0.0, similarity=20.497952]
37: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.6716046,
final_local_scores=0.668862, pem=0.0, similarity=24.049889]
38: 47905 Kolkata
    [pem_pos=0, final_scores=0.41847295,
final_local_scores=0.39585102, pem=-0.02273471,
similarity=16.026333]
39: 345921 Missionaries_of_Charity
    [pem_pos=0, final_scores=0.61051965,
final_local_scores=0.66129786, pem=0.0, similarity=23.82416]
40: 606848 Catholic_Church
    [pem_pos=0, final_scores=0.4527074,
final_local_scores=0.44864577, pem=0.0, similarity=17.478197]
41: 14533 India
    [pem_pos=0, final_scores=0.48175138,
final_local_scores=0.47652674, pem=-0.13622136,
similarity=19.051014]
42: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.4867447,
final_local_scores=0.491598, pem=-0.48936236,
```

```
similarity=21.421207]
43: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.5451471,
final_local_scores=0.53072643, pem=0.0, similarity=19.927649]
44: 606848 Catholic_Church
    [pem_pos=0, final_scores=0.2185878,
final_local_scores=0.19290566, pem=-0.38803694,
similarity=11.956608]
1: 236912 Bharat_Ratna
    [pem_pos=0, final_scores=0.07021187,
final_local_scores=0.07373841, pem=0.0, similarity=6.290214]
```

Sample: b'116_RTRS'

RTRS - FOCUS-News forecasts alien-led profit boost . Bernard Hickey [1 1 SYDNEY 1] 1996-08-22 Media baron [2 2 Rupert Murdoch 2] 's News Corp Ltd reported lower than expected 1995/96 profits on Thursday , but forecast that the hit film " 1 Independence Day " would help increase profits by at least 20 percent in 1996/97 . " From an earnings perspective , the current fiscal year has begun with great promise due to the hit motion picture ' 2 Independence Day , ' " [3 3 News Corp 3] said in a statement announcing its results for the year to June 30 , 1996 . It said moderating paper prices and solid orders for advertising at its [4 4 Fox Broadcasting 4] television network in the [5 5 United States 5] would also help boost profits in the 1996/97 year . " A budgeted profit increase of at least 20 percent for the full year currently appears very attainable , " [6 6 News Corp 6] said . The bullish comments for the coming year soothed analysts and most shareholders , who were disappointed by the lower than expected profit for 1995/96 . News announced pre-abnormals net profit for the year fell six percent to A\$ 1.26 billion ([7 7 US\$ 7] 995 million) and earnings per share dropped to 40 cents from 46 cents . Analysts had on average expected a pre-abnormals profit of A\$ 1.343 billion . " The year just gone was disappointing , but the outlook for the current year looks good , " [8 8 First Pacific 8] media analyst Lachlan Drummond said . [9 9 News Corp 9] said strong performances in [10 10 U.S. 10] television and [11 11 British 11] newspapers were offset by lower profits from [12 12 News Corp 12] 's magazine and publishing divisions and further hefty losses from its Asian Star TV operations . Higher newsprint prices hit profits hard . " Throughout the group , higher paper prices increased costs by over [13 13 US\$ 13] 300 million , " it said . [14 14 News Corp 14] said [15 15 British 15] newspaper operating profits rose 10 percent for the year , as higher cover prices at [16 16 The Sun 16] and [17 17 The Times 17] and higher advertising volumes offset increased newsprint costs . Advertising revenues at [18 18 The Times 18] grew 20 percent . Analysts said sharply lower earnings from [19 19 News Corp 19] 's book publishing division and its [20 20 U.S. 20] magazines had been the major surprises in the results for

1995/96 . [21 21 News Corp 21] said revenue gains at its magazines and inserts division were offset by higher paper prices and lower sales at the [22 22 U.S. 22] [23 23 TV Guide 23] . News said dramatically lower earnings from the [24 24 British 24] arm of its [25 25 Harper-Collins 25] publishing division more than offset healthy results from its [26 26 U.S. 26] operation . It said the demise of the Net Book Agreement had hurt the [27 27 British 27] operations , and weak performances from the [28 28 San Francisco 28] unit of [29 29 Harper-Collins 29] had not helped . The minimum price setting expired last September when three publishers pulled out . But it was the bullish profit forecast for 1996/97 that took the spotlight in the market , with some analysts saying 20 percent may even be an understatement . " If they 're saying at least 20 percent , then their internal forecasts are probably saying 25 or 30 percent , " said one [30 30 Sydney 30] media analyst who declined to be named . [31 31 News Corp 31] 's shares were down eight cents at A\$ 6.39 at 2.00 p.m. (0400 [32 32 GMT 32]) in a soft market . (A\$ 1 = [33 33 US\$ 33] 0.79) -- Sydney Newsroom 61-2 373-1800

1: 27862 Sydney
2: 26091 Rupert_Murdoch
3: 113422 News_Corporation
4: 46252 Fox_Broadcasting_Company
5: 3434750 United_States
6: 113422 News_Corporation
7: 18717338 United_States_dollar
8: 21902133 First_Pacific
9: 113422 News_Corporation
10: 3434750 United_States
11: 31717 United_Kingdom
12: 113422 News_Corporation
13: 18717338 United_States_dollar
14: 113422 News_Corporation
15: 31717 United_Kingdom
16: 26146987 The_Sun_(United_Kingdom)
17: 39127 The_Times
18: 39127 The_Times
19: 113422 News_Corporation
20: 3434750 United_States
21: 113422 News_Corporation
22: 3434750 United_States
23: 188834 TV_Guide
24: 31717 United_Kingdom
25: 253375 HarperCollins
26: 3434750 United_States
27: 31717 United_Kingdom
28: 49728 San_Francisco
29: 253375 HarperCollins
30: 27862 Sydney
31: 113422 News_Corporation
32: 12701 Greenwich_Mean_Time

```
33: 18717338 United_States_dollar
1: 52389 Independence_Day_(1996_film)
    [winner: 781729 Independence_Day_(United_States),
pem_pos=0, final_scores=-0.1313927,
final_local_scores=-0.13714676, pem=-0.3347267,
similarity=1.817276]
    RECALL MISS: 52389 Independence_Day_(1996_film)]
2: 52389 Independence_Day_(1996_film)
    [winner: 781729 Independence_Day_(United_States),
pem_pos=0, final_scores=-0.16376689,
final_local_scores=-0.17251064, pem=-0.3347267,
similarity=0.7619473]
    RECALL MISS: 52389 Independence_Day_(1996_film)]
1: 27862 Sydney
    [pem_pos=0, final_scores=0.61469,
final_local_scores=0.640509, pem=-0.1443489,
similarity=23.988771]
2: 26091 Rupert_Murdoch
    [pem_pos=0, final_scores=0.552763,
final_local_scores=0.5497085, pem=0.0, similarity=20.49411]
3: 113422 News_Corporation
    [pem_pos=0, final_scores=0.4332897,
final_local_scores=0.41488308, pem=0.0, similarity=16.470652]
4: 46252 Fox_Broadcasting_Company
    [pem_pos=0, final_scores=0.4975894,
final_local_scores=0.50771445, pem=0.0, similarity=19.240925]
5: 3434750 United_States
    [pem_pos=0, final_scores=0.5467782,
final_local_scores=0.5362559, pem=-0.073457584,
similarity=20.492132]
6: 113422 News_Corporation
    [pem_pos=0, final_scores=0.4305548,
final_local_scores=0.41189557, pem=0.0, similarity=16.381498]
7: 18717338 United_States_dollar
    [pem_pos=0, final_scores=0.50714266,
final_local_scores=0.5339175, pem=0.0, similarity=20.022875]
8: 21902133 First_Pacific
    [pem_pos=0, final_scores=0.45333904,
final_local_scores=0.5880537, pem=0.0, similarity=21.638409]
9: 113422 News_Corporation
    [pem_pos=0, final_scores=0.46081543,
final_local_scores=0.44495076, pem=0.0, similarity=17.367931]
10: 3434750 United_States
    [pem_pos=0, final_scores=0.547416,
final_local_scores=0.53695256, pem=-0.14075686,
similarity=20.878906]
11: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5656411,
final_local_scores=0.5512967, pem=-0.48936236,
similarity=23.202736]
12: 113422 News_Corporation
    [pem_pos=0, final_scores=0.5141178,
```

```
final_local_scores=0.5031755, pem=0.0, similarity=19.105473]
13: 18717338 United_States_dollar
    [pem_pos=0, final_scores=0.49333566,
final_local_scores=0.5188354, pem=0.0, similarity=19.572798]
14: 113422 News_Corporation
    [pem_pos=0, final_scores=0.43768722,
final_local_scores=0.41968673, pem=0.0, similarity=16.614002]
15: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5611102,
final_local_scores=0.5463474, pem=-0.48936236,
similarity=23.055037]
16: 26146987 The_Sun_(United_Kingdom)
    [pem_pos=0, final_scores=0.43755066,
final_local_scores=0.46438646, pem=-0.53386396,
similarity=20.851166]
17: 39127 The_Times
    [pem_pos=0, final_scores=0.4309265,
final_local_scores=0.43541217, pem=0.0, similarity=17.08328]
18: 39127 The_Times
    [pem_pos=0, final_scores=0.47325712,
final_local_scores=0.48165202, pem=0.0, similarity=18.46317]
19: 113422 News_Corporation
    [pem_pos=0, final_scores=0.50034195,
final_local_scores=0.48812747, pem=0.0, similarity=18.65641]
20: 3434750 United_States
    [pem_pos=0, final_scores=0.50883275,
final_local_scores=0.49480623, pem=-0.14075686,
similarity=19.621174]
21: 113422 News_Corporation
    [pem_pos=0, final_scores=0.48826307,
final_local_scores=0.4749331, pem=0.0, similarity=18.262663]
22: 3434750 United_States
    [pem_pos=0, final_scores=0.53517956,
final_local_scores=0.5235861, pem=-0.14075686,
similarity=20.480024]
23: 188834 TV_Guide
    [pem_pos=0, final_scores=0.49371803,
final_local_scores=0.55938864, pem=0.0, similarity=20.782986]
24: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.56363463,
final_local_scores=0.5491049, pem=-0.48936236,
similarity=23.137327]
25: 253375 HarperCollins
    [pem_pos=0, final_scores=0.3422926,
final_local_scores=0.3444182, pem=0.0, similarity=14.367839]
26: 3434750 United_States
    [pem_pos=0, final_scores=0.4694187,
final_local_scores=0.45175236, pem=-0.14075686,
similarity=18.33636]
27: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5374412,
final_local_scores=0.52049255, pem=-0.48936236,
```

```
similarity=22.283478]
28: 49728 San_Francisco
    [pem_pos=0, final_scores=0.45472735,
final_local_scores=0.45814103, pem=-0.09025385,
similarity=18.25237]
29: 253375 HarperCollins
    [pem_pos=0, final_scores=0.40184128,
final_local_scores=0.4094661, pem=0.0, similarity=16.308998]
30: 27862 Sydney
    [pem_pos=0, final_scores=0.4491822,
final_local_scores=0.45971692, pem=-0.1443489,
similarity=18.593575]
31: 113422 News_Corporation
    [pem_pos=0, final_scores=0.41757625,
final_local_scores=0.39771855, pem=0.0, similarity=15.958427]
32: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.66787225,
final_local_scores=0.76648676, pem=0.0, similarity=26.963207]
33: 18717338 United_States_dollar
    [pem_pos=0, final_scores=0.52067786,
final_local_scores=0.5487026, pem=0.0, similarity=20.464092]
```

Sample: b'117_RTRS'

RTRS - Budget cuts to boost [1 1 Australia 1] savings - RBA . [2 2 CANBERRA 2] 1996-08-22 The [3 3 Australian 3] government 's plans to slash its budget deficit should make a useful contribution to national savings , the [4 4 Reserve Bank of Australia 4] (RBA) said in its annual report . " The government 's announced plans to balance the budget , if realised , would make a useful contribution to raising national savings , " the RBA said . The bank said there were concerns fiscal consolidation would unduly restrict growth , but evidence was ambiguous . In its 1996/97 budget announced on Tuesday , the [5 5 Australian 5] [6 6 Coalition 6] government announced an underlying budget deficit of A\$ 5.65 billion , and pledged to return the underlying budget balance to surplus by 1998/99 . The budget deficit was A\$ 10.3 billion in 1995/96 . -- Sydney Newsroom 61-2 9373-1800 " Determined and credible efforts to rein in unsustainable fiscal positions (are) often rewarded by rising confidence , giving favourable effects on economic activity even in the short term , " it said . " More generally , the long-term effects of fiscal consolidation are clearly positive , with higher saving tending to promote economic growth by raising investment and lowering long-term real interest rates , " the RBA said .

```
1: 4689264 Australia
2: 51983 Canberra
3: 4689264 Australia
4: 492355 Reserve_Bank_of_Australia
5: 4689264 Australia
6: 1051313 Coalition_(Australia)
```

```
1: 4689264 Australia
    [pem_pos=0, final_scores=0.56196475,
final_local_scores=0.52431893, pem=-0.17168477,
similarity=20.670084]
2: 51983 Canberra
    [pem_pos=0, final_scores=0.49613816,
final_local_scores=0.47963202, pem=-0.091910355,
similarity=18.902712]
3: 4689264 Australia
    [pem_pos=0, final_scores=0.6325833,
final_local_scores=0.601459, pem=-0.491009,
similarity=24.708633]
4: 492355 Reserve_Bank_of_Australia
    [pem_pos=0, final_scores=0.48958606,
final_local_scores=0.5447051, pem=0.0, similarity=20.344799]
5: 4689264 Australia
    [pem_pos=0, final_scores=0.5942829,
final_local_scores=0.5596216, pem=-0.491009,
similarity=23.46012]
6: 1051313 Coalition_(Australia)
    [pem_pos=0, final_scores=0.43435556,
final_local_scores=0.5174947, pem=-0.54157776,
similarity=22.47797]
```

Sample: b'118_BNZ'

```
[1 1 BNZ 1] cuts NZ fixed home lending rates . [2 2 WELLINGT
ON 2] 1996-08-22 [3 3 Bank of New Zealand 3] said on Thursday
it was cutting its fixed home lending rates . The rates are :
New rate old rate Six month rate 10.5 pct 10.75 One year 10.5
pct 10.95 Two year 10.5 pct 11.25 Three year 10.5 pct 11.25 [4
4 BNZ 4] said it was responding to lower wholesale rates .
Fixed business and farm lending rates rates were left
unchanged although they were under review . -- [5 5 Wellington
5] newsroom 64 4 4734 746
```

```
1: 765639 Bank_of_New_Zealand
```

```
2: 33804 Wellington
```

```
3: 765639 Bank_of_New_Zealand
```

```
4: 765639 Bank_of_New_Zealand
```

```
5: 33804 Wellington
```

```
1: 765639 Bank_of_New_Zealand
```

```
    [pem_pos=0, final_scores=0.6394643,
final_local_scores=0.60912657, pem=0.0, similarity=22.267265]
```

```
2: 33804 Wellington
```

```
    [pem_pos=0, final_scores=0.43667418,
final_local_scores=0.4370579, pem=-0.9265605,
similarity=22.171173]
```

```
3: 765639 Bank_of_New_Zealand
```

```
    [pem_pos=0, final_scores=0.5353218,
final_local_scores=0.4953667, pem=0.0, similarity=18.872442]
```

```
4: 765639 Bank_of_New_Zealand
```

```
    [pem_pos=0, final_scores=0.64264804,
final_local_scores=0.6126044, pem=0.0, similarity=22.37105]
```

5: 33804 Wellington
[pem_pos=0, final_scores=0.49067026,
final_local_scores=0.49604046, pem=-0.9265605,
similarity=23.93133]

Sample: b'119_Power'
Power NZ ODV up 8 pct at [1 1 NZ\$ 1] 524 million . [2 2 WELL
INGTON 2] 1996-08-22 Power New Zealand said on Thursday that
the Optimised Deprival Value (ODV) of its network at March
31 , 1996 has been set at \$ 524.2 million , an increase of
eight percent on its \$ 486.5 million valuation a year
earlier . The company said the increase reflected the value
of extensions to the network to meet economic growth in its
supply area and an increase in the estimated lifespan of the
network . It said the increase was consistent with the
approach followed by other power companies and reflected the
company 's new levels of preventative maintenance and
equipment upgrading . The revaluation was undertaken to meet
the disclosure requirements of the Ministry of Commerce . -- [
3 3 Wellington 3] newsroom 64 4 4734 746

1: 161585 New_Zealand_dollar
2: 33804 Wellington
3: 33804 Wellington
1: 161585 New_Zealand_dollar
[pem_pos=0, final_scores=0.5488728,
final_local_scores=0.62274355, pem=0.0, similarity=22.673622]
2: 33804 Wellington
[pem_pos=0, final_scores=0.49550956,
final_local_scores=0.45297235, pem=-0.9265605,
similarity=22.646091]
3: 33804 Wellington
[pem_pos=0, final_scores=0.55363935,
final_local_scores=0.5164703, pem=-0.9265605,
similarity=24.540997]

Sample: b'120_Thais'
Thais hunt for [1 1 Australian 1] jail breaker . BANGKOK
Thailand has launched a manhunt for an [2 2 Australian 2] who
escaped from a high security prison in [3 3 Bangkok 3] while
awaiting trial on drug possession charges , officials said on
Thursday . Daniel Westlake , 46 , from [4 4 Victoria 4] ,
made the first sucessful escape from Klongprem prison in the
northern outskirts of the capital on Sunday night . He was
believed by prison officials to still be in [5 5 Thailand
5] . " We have ordered a massive hunt for him and I am quite
confident we will get him soon , " Vivit Chatuparisut ,
deputy director general of the Correction Department , told [6
6 Reuters 6] . Westlake , arrested in December 1993 and
charged with heroin trafficking , sawed the iron grill off
his cell window and climbed down the prison 's five-metre
(15-foot) wall on a rope made from bed sheets , Vivit
said . The corrections department was probing the escape and

had ordered all foreign inmates chained to prevent more breakouts . There are 266 Westerners , including six [7 7 Australians 7] , in the prison , most awaiting trial on drugs charges . There also are about 5,000 [8 8 Thai 8] inmates in Klongprem , a prison official said .

1: 4689264 Australia
2: 4689264 Australia
3: 56615 Bangkok
4: 4689460 Victoria_(Australia)
5: 30128 Thailand
6: 18998750 Reuters
7: 11197621 Australians
8: 30128 Thailand
1: 4689264 Australia
 [pem_pos=0, final_scores=0.75377625,
final_local_scores=0.72638285, pem=-0.491009,
similarity=28.43661]
2: 4689264 Australia
 [pem_pos=0, final_scores=0.60394436,
final_local_scores=0.56271434, pem=-0.491009,
similarity=23.552414]
3: 56615 Bangkok
 [pem_pos=0, final_scores=0.52803916,
final_local_scores=0.54268634, pem=-0.0005963434,
similarity=20.287798]
4: 4689460 Victoria_(Australia)
 [pem_pos=0, final_scores=0.3723336,
final_local_scores=0.36564976, pem=-0.6703371,
similarity=18.646828]
5: 30128 Thailand
 [pem_pos=0, final_scores=0.61210716,
final_local_scores=0.5953696, pem=-0.24029273,
similarity=23.163477]
6: 18998750 Reuters
 [pem_pos=0, final_scores=0.3209309,
final_local_scores=0.39764595, pem=0.0, similarity=15.956261]
7: 11197621 Australians
 [pem_pos=0, final_scores=0.2217197,
final_local_scores=0.21552235, pem=-0.7024679,
similarity=14.341459]
8: 30128 Thailand
 [pem_pos=2, final_scores=0.60519636,
final_local_scores=0.5878206, pem=-1.7504287,
similarity=31.150555]

Sample: b'121_Tokyo'

Tokyo Soir - 1996 parent forecast . [1 1 TOKYO 1] 1996-08-22
Year to December 31 , 1996 (in billions of yen unless
specified) LATEST ACTUAL (Parent) FORECAST YEAR-AGO Sales
26.00 26.70 Current 400 million 329 million Net 250 million
84 million EPS 11.61 yen 3.92 yen Ord div 10.00 yen 10.00 yen
NOTE - Tokyo Soir Co Ltd is a specialised manufacturer of

women " s formal wear .

1: 30057 Tokyo

1: 30057 Tokyo

[pem_pos=0, final_scores=0.4355651,
final_local_scores=0.7483759, pem=-0.09283106,
similarity=26.927572]

Sample: b'122_Ka'

[1 1 Ka Wah Bank 1] sets [2 2 HK\$ 2] 43 mln FRCD . [3 3 HONG KONG 3] 1996-08-22 [4 4 Ka Wah Bank 4] 's [5 5 HK\$ 5] 43 million floating rate certificate of deposit issue has been privately placed , sole arranger HSBC Markets said . The facility has a tenor of six months . It pays a coupon of 15 basis points over the six-month Hong Kong Interbank Offered Rate . Other details are not available . The deposit date is September 5 , 1996 . Clearing is through the Hong Kong Central Moneymarkets Unit . -- Hong Kong Newsroom (852) 2847 4039

1: 22014010 Ka_Wah_Bank

2: 309734 Hong_Kong_dollar

3: 13404 Hong_Kong

4: 22014010 Ka_Wah_Bank

5: 309734 Hong_Kong_dollar

1: 22014010 Ka_Wah_Bank

[pem_pos=0, final_scores=0.5296022,
final_local_scores=0.60827506, pem=0.0, similarity=22.241854]

2: 309734 Hong_Kong_dollar

[pem_pos=0, final_scores=0.6240872,
final_local_scores=0.6577836, pem=0.0, similarity=23.719288]

3: 13404 Hong_Kong

[pem_pos=0, final_scores=0.49270874,
final_local_scores=0.5535632, pem=-0.22402541,
similarity=21.827427]

4: 22014010 Ka_Wah_Bank

[pem_pos=0, final_scores=0.5031454,
final_local_scores=0.57937497, pem=0.0, similarity=21.379417]

5: 309734 Hong_Kong_dollar

[pem_pos=0, final_scores=0.5525209,
final_local_scores=0.57960826, pem=0.0, similarity=21.38638]

Sample: b'123_Malaysia'

[1 1 Malaysia 1] bans nitrofuran usage in chicken feed . [2 2 KUALA LUMPUR 2] 1996-08-22 [3 3 Malaysia 3] has banned the use of nitrofuran , an antibiotic , in chicken feed and veterinary applications because it believes the drug could cause cancer , the health ministry said on Thursday . " It is hoped that livestock breeders and feedmillers will abide by the laws and respect the cabinet decision in the interest of consumer safety , " Health Minister Chua Jui Meng was quoted as saying by the national [4 4 Bernama 4] news agency . [5 5 C hua 5] said offenders could face a two-year prison sentence and a maximum fine of 5,000 ringgit (\$ 2000) . " The ban

takes effect immediately , " he added .

1: 3607937 Malaysia
2: 16854 Kuala_Lumpur
3: 3607937 Malaysia
4: 1143836 Bernama
5: 16429740 Chua_Jui_Meng
1: 3607937 Malaysia
 [pem_pos=0, final_scores=0.65479916,
final_local_scores=0.65824765, pem=-0.20223404,
similarity=24.832916]
2: 16854 Kuala_Lumpur
 [pem_pos=0, final_scores=0.43269944,
final_local_scores=0.43574232, pem=0.0, similarity=17.093132]
3: 3607937 Malaysia
 [pem_pos=0, final_scores=0.5658029,
final_local_scores=0.5610327, pem=-0.20223404,
similarity=21.931828]
4: 1143836 Bernama
 [pem_pos=0, final_scores=0.6068026,
final_local_scores=0.6789108, pem=0.0, similarity=24.349766]
5: 16429740 Chua_Jui_Meng
 [pem_pos=0, final_scores=0.94164807,
final_local_scores=1.0703266, pem=0.0, similarity=36.030396]

Sample: b'124_INDONESIAN'

[1 1 INDONESIAN 1] STOCKS - factors to watch - August 22 . [2 2 JAKARTA 2] 1996-08-22 Following are some of the main factors likely to affect [3 3 Indonesian 3] stocks on Thursday : ** Security was tight in [4 4 Jakarta 4] ahead of a trial involving ousted [5 5 Indonesian Democratic Party 5] leader [6 6 Megawati Sukarnoputri 6] . Around 200 police and troops were stationed outside the court in central [7 7 Jakarta 7] but there was no sign of demonstrators . ** The [8 8 Dow Jones 8] industrial average closed down 31.44 points at 5,689.82 on Wednesday , ending a three-session winning streak as investors took profits and tobacco stocks took a beating . MARKETS : ** The [9 9 Jakarta 9] composite index rose 2.60 points , or 0.48 percent , to 542.20 points on Wednesday on the back of bargain-hunting in selected big-capitalised stocks and secondliners . ** On Thursday , the [10 10 Indonesian 10] rupiah was at 2,343.00 / 43.50 in early trading against an opening of 2,342.75 / 43.50 . STOCKS TO WATCH ** Packaging manufacturer Super Indah Makmur on announcement of a tender offer by PT VDH Teguh Sakti , a wholly-owned subsidiary of Singapore-listed Van Der Horst . ** Privately-owned Bank Duta on market talk that it is obtaining fresh syndicated loans , a management reshuffle and fresh equity injection . ** Ciputra Development on reports of a plan to build property projects worth \$ 2 billion in [11 11 Jakarta 11] and [12 12 Surabaya 12] .

1: 14579 Indonesia
2: 16275 Jakarta

```
3: 14579 Indonesia
4: 16275 Jakarta
5: 1010736 Indonesian_Democratic_Party
6: 237564 Megawati_Sukarnoputri
7: 16275 Jakarta
8: 47361 Dow_Jones_Industrial_Average
9: 16275 Jakarta
10: 14579 Indonesia
11: 16275 Jakarta
12: 212635 Surabaya
1: 14579 Indonesia
    [pem_pos=1, final_scores=0.46602052,
final_local_scores=0.43065524, pem=-0.74208045,
similarity=20.976873]
2: 16275 Jakarta
    [pem_pos=0, final_scores=0.46437055,
final_local_scores=0.3970422, pem=0.0, similarity=15.938244]
3: 14579 Indonesia
    [pem_pos=1, final_scores=0.4585839,
final_local_scores=0.42253184, pem=-0.74208045,
similarity=20.734455]
4: 16275 Jakarta
    [pem_pos=0, final_scores=0.48251003,
final_local_scores=0.41685688, pem=0.0, similarity=16.529554]
5: 1010736 Indonesian_Democratic_Party
    [pem_pos=0, final_scores=0.42756432,
final_local_scores=0.53543895, pem=0.0, similarity=20.06828]
6: 237564 Megawati_Sukarnoputri
    [pem_pos=0, final_scores=0.56310934,
final_local_scores=0.6139423, pem=0.0, similarity=22.410976]
7: 16275 Jakarta
    [pem_pos=0, final_scores=0.45810968,
final_local_scores=0.39020324, pem=0.0, similarity=15.734156]
8: 47361 Dow_Jones_Industrial_Average
    [pem_pos=1, final_scores=0.29692745,
final_local_scores=0.44094634, pem=-0.42373553,
similarity=19.55277]
9: 16275 Jakarta
    [pem_pos=0, final_scores=0.49930066,
final_local_scores=0.43519813, pem=0.0, similarity=17.076893]
10: 14579 Indonesia
    [pem_pos=1, final_scores=0.4982801,
final_local_scores=0.46589398, pem=-0.74208045,
similarity=22.028467]
11: 16275 Jakarta
    [pem_pos=0, final_scores=0.4286496,
final_local_scores=0.3580226, pem=0.0, similarity=14.773822]
12: 212635 Surabaya
    [pem_pos=0, final_scores=0.46681708,
final_local_scores=0.43473947, pem=0.0, similarity=17.063206]
```

Sample: b'125_Key'

Key stock and currency market movements at 1600 [1 1 GMT 1] .
[2 2 LONDON 2] 1996-08-23 The following table shows the latest close of key indices* on major world stock exchanges , the day 's change in points and the indices ' 1996 closing highs and lows (with dates) . Also shown are the [3 3 London 3] closing values of the [4 4 German 4] mark , the [5 5 Japan ese 5] yen , the [6 6 British 6] pound and gold bullion (previous day 's closes in brackets) : AUG 23 DAY 'S CHANGE 1996 HIGH 1996 LOW CLOSE IN POINTS [7 7 NEW YORK 7] 5,710.53 - 22.94 5,778.00 5,032.94 (midday) (May 22) (Jan 10) [8 8 LONDON 8] 3,907.5 +16.4 3,907.5 3,632.3 (Aug 23) (Jul 16) [9 9 TOKYO 9] 21,228.80 - 134.44 22,666.80 19,734.70 (Jun 26) (Mar 13) [10 10 FRANKFURT 10] 2,555.16 - 2.10 2,583.49) 2,284.86 (Jul 5) (Jan 2) [11 11 PARIS 11] 2,020.82 +3.06 2,146.79 1,897.85 (Apr 30) (Jan 11) [12 12 SYDNEY 12] 2,292.9 +18.3 2,326.00 2,096.10 (Apr 26) (Jul 17) [13 13 HONG KONG 13] 11,424.64 - 54.13 11,594.99 10,204.87 (Feb 16) (Jan 2) - - - - FOREIGN EXCHANGE / GOLD BULLION CLOSE IN [14 14 LONDON 14] Dollar / mark ... 1.4871 (1.4935) Dollar / yen 108.50 (108.43) Pound / dollar .. \$ 1.5520 (\$ 1.5497) Gold (ounce) .. \$ 387.50 (\$ 386.95) - - - - *INDICES USED AND THEIR ALL-TIME CLOSING HIGHS [15 15 New York 15] [16 16 Dow Jones 16] industrial average -- 5,778.00 (May 22/96) [17 17 London 17] [18 18 FT SE-100 18] index -- 3,907.5 (Aug 23/96) [19 19 Tokyo 19] [20 20 Nikkei 20] average -- 38,915.87 (Dec 29/89) [21 21 Frank furt 21] DAX-30 index -- 2,583.49 (Jul 5/96) [22 22 Paris 22] CAC-40 General index -- 2,355.93 (Feb 2/94) [23 23 Sydney 23] Australian All-Ordinaries index -- 2,340.6 (Feb 3/94) [2 4 24 Hong Kong 24] [25 25 Hang Seng 25] index -- 12,201.09 (Jan 4/94)

1: 12701 Greenwich_Mean_Time
2: 17867 London
3: 17867 London
4: 11867 Germany
5: 15573 Japan
6: 31717 United_Kingdom
7: 645042 New_York_City
8: 17867 London
9: 30057 Tokyo
10: 10992 Frankfurt
11: 22989 Paris
12: 27862 Sydney
13: 13404 Hong_Kong
14: 17867 London
15: 645042 New_York_City
16: 47361 Dow_Jones_Industrial_Average
17: 17867 London
18: 497717 FTSE_100_Index
19: 30057 Tokyo
20: 40293 Nikkei_225
21: 10992 Frankfurt

```
22: 22989 Paris
23: 27862 Sydney
24: 13404 Hong_Kong
25: 23476353 Hang_Seng_Index
1: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.61262745,
final_local_scores=0.70266265, pem=0.0, similarity=25.058567]
2: 17867 London
    [pem_pos=0, final_scores=0.7696381,
final_local_scores=0.73960906, pem=-0.09086737,
similarity=26.655273]
3: 17867 London
    [pem_pos=0, final_scores=0.5710861,
final_local_scores=0.5227212, pem=-0.09086737,
similarity=20.182905]
4: 11867 Germany
    [pem_pos=1, final_scores=0.7832354,
final_local_scores=0.80917925, pem=-1.0465823,
similarity=33.928715]
5: 15573 Japan
    [pem_pos=0, final_scores=0.73419195,
final_local_scores=0.77116126, pem=-0.99076414,
similarity=32.49063]
6: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.52677554,
final_local_scores=0.5030865, pem=-0.48936236,
similarity=21.764046]
7: 645042 New_York_City
    [pem_pos=1, final_scores=0.46619302,
final_local_scores=0.4326877, pem=-1.7906687,
similarity=26.739912]
8: 17867 London
    [pem_pos=0, final_scores=0.63314146,
final_local_scores=0.59050715, pem=-0.09086737,
similarity=22.205774]
9: 30057 Tokyo
    [pem_pos=0, final_scores=0.7299224,
final_local_scores=0.7310097, pem=-0.09283106,
similarity=26.40933]
10: 10992 Frankfurt
    [pem_pos=0, final_scores=0.44527543,
final_local_scores=0.42950243, pem=-0.07062824,
similarity=17.29101]
11: 22989 Paris
    [pem_pos=0, final_scores=0.4083283,
final_local_scores=0.3894874, pem=-0.115301244,
similarity=16.339819]
12: 27862 Sydney
    [pem_pos=0, final_scores=0.5934079,
final_local_scores=0.5776398, pem=-0.1443489,
similarity=22.112629]
13: 13404 Hong_Kong
```

```
[pem_pos=0, final_scores=0.52008814,
final_local_scores=0.48425162, pem=-0.22402541,
similarity=19.759033]
14: 17867 London
    [pem_pos=0, final_scores=0.5659208,
final_local_scores=0.5170789, pem=-0.09086737,
similarity=20.014528]
15: 645042 New_York_City
    [pem_pos=1, final_scores=0.43849343,
final_local_scores=0.40243018, pem=-1.7906687,
similarity=25.836967]
16: 47361 Dow_Jones_Industrial_Average
    [pem_pos=1, final_scores=0.35625857,
final_local_scores=0.4400416, pem=-0.42373553,
similarity=19.52577]
17: 17867 London
    [pem_pos=0, final_scores=0.6397323,
final_local_scores=0.5977067, pem=-0.09086737,
similarity=22.420622]
18: 497717 FTSE_100_Index
    [pem_pos=0, final_scores=0.5484438,
final_local_scores=0.62514806, pem=0.0, similarity=22.745378]
19: 30057 Tokyo
    [pem_pos=0, final_scores=0.6268909,
final_local_scores=0.6184634, pem=-0.09283106,
similarity=23.050724]
20: 40293 Nikkei_225
    [pem_pos=1, final_scores=0.31909543,
final_local_scores=0.43435216, pem=-0.46501607,
similarity=19.580477]
21: 10992 Frankfurt
    [pem_pos=0, final_scores=0.44361812,
final_local_scores=0.42769212, pem=-0.07062824,
similarity=17.236986]
22: 22989 Paris
    [pem_pos=0, final_scores=0.65689677,
final_local_scores=0.66101074, pem=-0.115301244,
similarity=24.442617]
23: 27862 Sydney
    [pem_pos=0, final_scores=0.7491554,
final_local_scores=0.74777025, pem=-0.1443489,
similarity=27.189661]
24: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5040621,
final_local_scores=0.46674556, pem=-0.22402541,
similarity=19.236616]
25: 23476353 Hang_Seng_Index
    [pem_pos=0, final_scores=0.46964735,
final_local_scores=0.59986717, pem=0.0, similarity=21.990946]
```

Sample: b'126_Ukraine'

[1 1 Ukraine 1] hails peace as marks five-year

independence . Rostislav Khotin [2 2 KIEV 2] 1996-08-23 [3 3 Ukraine 3] celebrates five years of independence from [4 4 Kremlin 4] rule on Saturday , hailing civil and inter-ethnic peace as its main post-Soviet achievement . [5 5 Ukraine 5] 's declaration of independence in 1991 , backed nine-to-one by a referendum in December of that year , effectively dealt a death blow to the [6 6 Soviet 6] empire and ended more than three centuries of rule from [7 7 Moscow 7] . [8 8 Ukraine 8] , with a [9 9 Russian 9] community of 11 million people -- the world 's largest outside [10 10 Russia 10] -- has avoided conflicts like those in [11 11 Russia 11] 's [12 12 Chechnya 12] , neighbouring [13 13 Moldova 13] , and the former [14 14 Soviet 14] republics of [15 15 Georgia 15] , [16 16 Azerbaijan 16] and [17 17 Tajikistan 17] . " [18 18 Ukraine 18] 's biggest achievements for five years are the preservation of civil peace and inter-ethnic harmony , " President [19 19 Leonid Kuchma 19] said in televised statement this week . " Unlike many other [20 20 post-Soviet 20] countries we were able to deal with conflict situations in a peaceful and civilised way . " But independence was initially accompanied by hyper-inflation and economic collapse , although there are signs of a turnaround . Inflation -- a hyper-inflationary 10,300 percent a year in 1993 -- was a respectable 0.1 percent a month in June and July and the economy has just begun to grow . [21 21 Kuchma 21] told a solemn ceremony at the Ukraina Palace on Friday that " there was a turning point " in reforms and that he expected a rise in the standard of living in the near future . " There is no doubt that economic growth has already started , " said Adelbert Knobl , head of the [22 22 International Monetary Fund 22] 's mission in [23 23 Ukraine 23] . " The national bank and the government have every reason to be proud of their efforts . " Central bank officials said on Thursday that a much-postponed hryvna currency would " definitely " be introduced before the end of this year . It will replace the interim karbovanets currency , which was introduced at par to the [24 24 Russian 24] rouble in 1992 but now trades at almost 33 karbovanets per rouble . [25 25 Ukraine 25] has repeatedly promised to introduce the hryvna but had to postpone the plans because of economic problems . Proud of its record in promptly joining both the [26 26 Council of Europe 26] and [27 27 NATO 27] 's Partnership for Peace , [28 28 Ukraine 28] caused a foreign policy wrangle this week , offending [29 29 China 29] by allowing a [30 30 Taiwanese 30] minister to appear on a public , if unofficial visit . [31 31 China 31] cancelled a visit by a top-level delegation in protest . Kiev 's Foreign Minister Hennady Udovenko said [32 32 Beijing 32] was overreacting . But [33 33 Ukraine 33] , seeing itself as a bridge between [34 34 Russia 34] and the rapidly Westernising countries of eastern [35 35 Europe 35] , is looking West as well as East . " The strategic aim of [36 36 European 36] integration should not in any way damage [37 37 Ukraine 37]

's interests in [38 38 post-Soviet 38] areas . Relations with [39 39 Russia 39] , which is our main partner , have great importance , " [40 40 Kuchma 40] said . " But [41 41 Ukraine 4 1] cannot be economically oriented on [42 42 Russia 42] , even though those in some circles push us to do that . " [43 4 3 Kuchma 43] has said Kiev wants membership of the [44 44 European Union 44] , associate membership of the Western European Union defence grouping and to move closer to [45 45 NATO 45] . A message from the West this week from [46 46 U.S. 4 6] President [47 47 Bill Clinton 47] congratulated [48 48 Ukraine 48] on the anniversary , promising to support market reforms and praising [49 49 Ukraine 49] as a " stabilising factor " in a united [50 50 Europe 50] .

- 1: 31750 Ukraine
- 2: 585629 Kiev
- 3: 31750 Ukraine
- 4: 9787373 Government_of_Russia
- 5: 31750 Ukraine
- 6: 26779 Soviet_Union
- 7: 19004 Moscow
- 8: 31750 Ukraine
- 9: 25391 Russia
- 10: 25391 Russia
- 11: 25391 Russia
- 12: 6095 Chechnya
- 13: 19260 Moldova
- 14: 26779 Soviet_Union
- 15: 48768 Georgia_(country)
- 16: 746 Azerbaijan
- 17: 30108 Tajikistan
- 18: 31750 Ukraine
- 19: 383797 Leonid_Kuchma
- 20: 1609192 Post-Soviet_states
- 21: 383797 Leonid_Kuchma
- 22: 15251 International_Monetary_Fund
- 23: 31750 Ukraine
- 24: 25391 Russia
- 25: 31750 Ukraine
- 26: 5865 Council_of_Europe
- 27: 21133 NATO
- 28: 31750 Ukraine
- 29: 5405 China
- 30: 25734 Taiwan
- 31: 5405 China
- 32: 5405 China
- 33: 31750 Ukraine
- 34: 25391 Russia
- 35: 9239 Europe
- 36: 9239 Europe
- 37: 31750 Ukraine
- 38: 1609192 Post-Soviet_states
- 39: 25391 Russia

40: 383797 Leonid_Kuchma
41: 31750 Ukraine
42: 25391 Russia
43: 383797 Leonid_Kuchma
44: 9317 European_Union
45: 21133 NATO
46: 3434750 United_States
47: 3356 Bill_Clinton
48: 31750 Ukraine
49: 31750 Ukraine
50: 9239 Europe
1: 31750 Ukraine
[pem_pos=0, final_scores=0.6299495,
final_local_scores=0.57998896, pem=-0.35407752,
similarity=23.323269]
2: 585629 Kiev
[pem_pos=0, final_scores=0.45404232,
final_local_scores=0.4222836, pem=-0.032242518,
similarity=16.866837]
3: 31750 Ukraine
[pem_pos=0, final_scores=0.6544765,
final_local_scores=0.60678095, pem=-0.35407752,
similarity=24.122795]
4: 9787373 Government_of_Russia
[pem_pos=3, final_scores=0.5640315,
final_local_scores=0.59439516, pem=-2.0052693,
similarity=32.732616]
5: 31750 Ukraine
[pem_pos=0, final_scores=0.71447057,
final_local_scores=0.67231536, pem=-0.35407752,
similarity=26.078472]
6: 26779 Soviet_Union
[pem_pos=0, final_scores=0.41647148,
final_local_scores=0.37693584, pem=-0.097171314,
similarity=15.866663]
7: 19004 Moscow
[pem_pos=0, final_scores=0.47410125,
final_local_scores=0.44527268, pem=-0.106582455,
similarity=17.95715]
8: 31750 Ukraine
[pem_pos=0, final_scores=0.63882166,
final_local_scores=0.58968043, pem=-0.35407752,
similarity=23.612482]
9: 25391 Russia
[pem_pos=1, final_scores=0.82288486,
final_local_scores=0.80039513, pem=-1.0515022,
similarity=33.693333]
10: 25391 Russia
[pem_pos=1, final_scores=0.7723178,
final_local_scores=0.74515826, pem=-0.20299491,
similarity=27.43064]
11: 25391 Russia

```
[pem_pos=1, final_scores=0.86485773,
final_local_scores=0.84624404, pem=-0.20299491,
similarity=30.44724]
12: 6095 Chechnya
    [pem_pos=0, final_scores=0.48588437,
final_local_scores=0.49359226, pem=0.0, similarity=18.81949]
13: 19260 Moldova
    [pem_pos=0, final_scores=0.407435,
final_local_scores=0.37762052, pem=-0.47343826,
similarity=17.933294]
14: 26779 Soviet_Union
    [pem_pos=0, final_scores=0.46933502,
final_local_scores=0.43468124, pem=-0.097171314,
similarity=17.589901]
15: 48768 Georgia_(country)
    [pem_pos=1, final_scores=0.46003467,
final_local_scores=0.4195937, pem=-1.3466994,
similarity=23.934786]
16: 746 Azerbaijan
    [pem_pos=0, final_scores=0.4173562,
final_local_scores=0.38407332, pem=-0.33647978,
similarity=17.381056]
17: 30108 Tajikistan
    [pem_pos=0, final_scores=0.7078911,
final_local_scores=0.72267985, pem=-0.36836085,
similarity=27.659124]
18: 31750 Ukraine
    [pem_pos=0, final_scores=0.6668236,
final_local_scores=0.6202683, pem=-0.35407752,
similarity=24.525284]
19: 383797 Leonid_Kuchma
    [pem_pos=0, final_scores=0.7980984,
final_local_scores=0.85483795, pem=0.0, similarity=29.599783]
20: 1609192 Post-Soviet_states
    [pem_pos=0, final_scores=0.48443633,
final_local_scores=0.46214908, pem=0.0, similarity=17.881163]
21: 383797 Leonid_Kuchma
    [pem_pos=0, final_scores=0.4197622,
final_local_scores=0.4415632, pem=0.0, similarity=17.26684]
22: 15251 International_Monetary_Fund
    [pem_pos=0, final_scores=0.42809814,
final_local_scores=0.50103825, pem=0.0, similarity=19.041695]
23: 31750 Ukraine
    [pem_pos=0, final_scores=0.68880004,
final_local_scores=0.6442742, pem=-0.35407752,
similarity=25.241667]
24: 25391 Russia
    [pem_pos=2, final_scores=0.70574313,
final_local_scores=0.6724355, pem=-1.0515022,
similarity=29.874762]
25: 31750 Ukraine
    [pem_pos=0, final_scores=0.585797,
```

```
final_local_scores=0.531759, pem=-0.35407752,
similarity=21.883991]
26: 5865 Council_of_Europe
    [pem_pos=0, final_scores=0.34059066,
final_local_scores=0.34822565, pem=0.0, similarity=14.481461]
27: 21133 NATO
    [pem_pos=0, final_scores=0.4434303,
final_local_scores=0.45790702, pem=0.0, similarity=17.754572]
28: 31750 Ukraine
    [pem_pos=0, final_scores=0.6794465,
final_local_scores=0.6340569, pem=-0.35407752,
similarity=24.936764]
29: 5405 China
    [pem_pos=0, final_scores=0.5397306,
final_local_scores=0.52911305, pem=-0.15669827,
similarity=20.731651]
30: 25734 Taiwan
    [pem_pos=0, final_scores=0.5579699,
final_local_scores=0.57136554, pem=-0.76474255,
similarity=25.299189]
31: 5405 China
    [pem_pos=0, final_scores=0.5257891,
final_local_scores=0.513884, pem=-0.15669827,
similarity=20.277187]
32: 5405 China
    [pem_pos=7, final_scores=0.31032318,
final_local_scores=0.2785203, pem=-3.4768412,
similarity=31.308916]
33: 31750 Ukraine
    [pem_pos=0, final_scores=0.6567609,
final_local_scores=0.6092763, pem=-0.35407752,
similarity=24.197262]
34: 25391 Russia
    [pem_pos=1, final_scores=0.8645374,
final_local_scores=0.84589416, pem=-0.20299491,
similarity=30.4368]
35: 9239 Europe
    [pem_pos=0, final_scores=0.620284,
final_local_scores=0.61860234, pem=-0.086338006,
similarity=23.01956]
36: 9239 Europe
    [pem_pos=0, final_scores=0.42883378,
final_local_scores=0.4094721, pem=-0.25557294,
similarity=17.699022]
37: 31750 Ukraine
    [pem_pos=0, final_scores=0.6513865,
final_local_scores=0.6034056, pem=-0.35407752,
similarity=24.022068]
38: 1609192 Post-Soviet_states
    [pem_pos=0, final_scores=0.44207096,
final_local_scores=0.41587132, pem=0.0, similarity=16.500143]
39: 25391 Russia
```

```
[pem_pos=0, final_scores=0.85392314,
final_local_scores=0.83429974, pem=-0.20299491,
similarity=30.0908]
40: 383797 Leonid_Kuchma
    [pem_pos=0, final_scores=0.45711106,
final_local_scores=0.4823612, pem=0.0, similarity=18.484333]
41: 31750 Ukraine
    [pem_pos=0, final_scores=0.6559361,
final_local_scores=0.6083754, pem=-0.35407752,
similarity=24.170376]
42: 25391 Russia
    [pem_pos=1, final_scores=0.80957764,
final_local_scores=0.785859, pem=-0.20299491,
similarity=28.64523]
43: 383797 Leonid_Kuchma
    [pem_pos=0, final_scores=0.51336867,
final_local_scores=0.543814, pem=0.0, similarity=20.318207]
44: 9317 European_Union
    [pem_pos=0, final_scores=0.5003963,
final_local_scores=0.5041855, pem=-0.005814796,
similarity=19.167236]
45: 21133 NATO
    [pem_pos=0, final_scores=0.4157219,
final_local_scores=0.42763978, pem=0.0, similarity=16.851337]
46: 3434750 United_States
    [pem_pos=0, final_scores=0.5122681,
final_local_scores=0.5120041, pem=-0.14075686,
similarity=20.134394]
47: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.50564367,
final_local_scores=0.58221257, pem=0.0, similarity=21.464098]
48: 31750 Ukraine
    [pem_pos=0, final_scores=0.59645003,
final_local_scores=0.5433959, pem=-0.35407752,
similarity=22.231258]
49: 31750 Ukraine
    [pem_pos=0, final_scores=0.6254546,
final_local_scores=0.57507896, pem=-0.35407752,
similarity=23.176744]
50: 9239 Europe
    [pem_pos=0, final_scores=0.59731627,
final_local_scores=0.59351355, pem=-0.086338006,
similarity=22.27086]
```

Sample: b'127_Oldest'

Oldest [1 1 Albania 1] book disappears from [2 2 Vatican 2]
- paper . [3 3 TIRANA 3] 1996-08-23 A 16th-century document ,
the earliest complete example of written [4 4 Albanian 4] ,
has disappeared from the [5 5 Vatican 5] archives , an [6 6 Al
banian 6] newspaper said on Friday . Gazeta Shqiptare said
the " Book of Mass ' , by [7 7 Gjon Buzuku 7] , dating from
1555 and discovered in 1740 in a religious seminary in [8 8 Ro

me 8] , was the first major document published in the [9 9 Albanian 9] language . " We [10 10 Albanians 10] , sons of [11 1 1 Buzuku 11] , believed our language had a written document but now we do not have it any more , " lamented scholar Musa Hamiti , told of the loss by the [12 12 Vatican 12] library . [13 13 Tirana 13] 's national library has three copies of the " Book of Mass ' . " There is nothing left for us but to be grateful to civilisation for inventing photocopies , " Gazeta Shqiptare said .

1: 738 Albania
2: 462415 Vatican_Library
3: 31075 Tirana
4: 18940550 Albanian_language
5: 462415 Vatican_Library
6: 738 Albania
7: 333292 Gjon_Buzuku
8: 25458 Rome
9: 18940550 Albanian_language
10: 2927 Albanians
11: 333292 Gjon_Buzuku
12: 462415 Vatican_Library
13: 31075 Tirana
1: 738 Albania
[pem_pos=0, final_scores=0.5091221,
final_local_scores=0.50237066, pem=-0.41162592,
similarity=21.31994]
2: 462415 Vatican_Library
[pem_pos=7, final_scores=0.5140806,
final_local_scores=0.5517928, pem=-2.7816074,
similarity=35.683125]
3: 31075 Tirana
[pem_pos=0, final_scores=0.40335447,
final_local_scores=0.38568568, pem=0.0, similarity=15.599343]
4: 18940550 Albanian_language
[pem_pos=0, final_scores=0.32240695,
final_local_scores=0.30455896, pem=-0.7457155,
similarity=17.233677]
5: 462415 Vatican_Library
[pem_pos=7, final_scores=0.44895142,
final_local_scores=0.48064905, pem=-2.7816074,
similarity=33.56005]
6: 738 Albania
[pem_pos=2, final_scores=0.49986845,
final_local_scores=0.49226254, pem=-1.5560378,
similarity=27.241785]
7: 333292 Gjon_Buzuku
[pem_pos=0, final_scores=0.48378497,
final_local_scores=0.54605705, pem=0.0, similarity=20.385145]
8: 25458 Rome
[pem_pos=0, final_scores=0.47869068,
final_local_scores=0.47857577, pem=-0.2517334,
similarity=19.740334]

```
9: 18940550 Albanian_language
    [pem_pos=0, final_scores=0.2912976,
final_local_scores=0.27057675, pem=-0.7457155,
similarity=16.21958]
10: 2927 Albanians
    [pem_pos=0, final_scores=0.33688223,
final_local_scores=0.322741, pem=-0.028324477,
similarity=13.874981]
11: 333292 Gjon_Buzuku
    [pem_pos=0, final_scores=0.544449,
final_local_scores=0.6123232, pem=0.0, similarity=22.36266]
12: 462415 Vatican_Library
    [pem_pos=7, final_scores=0.3824154,
final_local_scores=0.40796858, pem=-2.7816074,
similarity=31.391123]
13: 31075 Tirana
    [pem_pos=0, final_scores=0.3475663,
final_local_scores=0.32474563, pem=0.0, similarity=13.78077]
```

Sample: b'128_Russia'

[1 1 Russia 1] to clamp down on barter deals . [2 2 MOSCOW 2]
1996-08-23 [3 3 Russian 3] officials , keen to cut capital flight , will adopt tight measures to cut barter deals in foreign trade to a minimum , a customs official said on Friday . " We have always been concerned about barter deals with other countries , viewing them as a disguised kind of capital flight from [4 4 Russia 4] , " Marina Volkova , deputy head of the currency department at the State Customs Committee , told [5 5 Reuters 5] . Volkova said last year goods had been exported under many [6 6 Russian 6] barter deals , with nothing imported in return . She said the cost of such unimported goods was \$ 1.10 billion in 1995 . Barter deals were worth \$ 4.9 billion last year , or about eight percent of all [7 7 Russian 7] exports estimated at \$ 61.5 billion , she said . " The cost of exported goods is too often understated , so the actual share of barter deals in [8 8 Russian 8] exports and the amount of unimported goods may be even higher , " Volkova said . A few days ago [9 9 Russian 9] President [10 10 Boris Yeltsin 10] issued a decree on state regulation of foreign barter deals , and Volkova said this " could substantially improve the situation " . In line with the decree , which will come into force on November 1 , all [11 11 Russian 11] barter traders will be obliged to import goods worth the cost of their exports within 180 days . " If traders are late , they will have to pay fines worth the cost of their exported goods , " Volkova said . Understating the cost of exported goods could still be a loophole for barter dealers , but Volkova said the authorities are currently " tackling the technicalities of the issue " . Barter has always been a feature of the [12 12 Soviet Union 12] 's foreign trade , but [13 13 Yeltsin 13] 's decrees liberalising foreign trade in 1991-1992 has given

barter a new impetus . A few years ago , barter deals accounted for up to 25-30 percent of [14 14 Russian 14] exports because " thousands (of) trade companies which popped up preferred barter in the absence of reliable [15 15 R ussian 15] banks and money transfer systems " , Volkova said . " Now many [16 16 Russian 16] banks are strong and can make various sorts of money transfers , while incompetent traders are being ousted by more experienced ones . But the current share of barter deals in [17 17 Russian 17] exports is still high , " she said . -- Dmitry Solovyov , Moscow Newsroom , +7095 941 8520

1: 25391 Russia
2: 19004 Moscow
3: 25391 Russia
4: 25391 Russia
5: 18998750 Reuters
6: 25391 Russia
7: 25391 Russia
8: 25391 Russia
9: 25391 Russia
10: 23278520 Boris_Yeltsin
11: 25391 Russia
12: 26779 Soviet_Union
13: 23278520 Boris_Yeltsin
14: 25391 Russia
15: 25391 Russia
16: 25391 Russia
17: 25391 Russia
1: 25391 Russia
[pem_pos=0, final_scores=0.7917016,
final_local_scores=0.74456507, pem=-0.20299491,
similarity=27.412937]
2: 19004 Moscow
[pem_pos=0, final_scores=0.66313356,
final_local_scores=0.65664643, pem=-0.106582455,
similarity=24.264965]
3: 25391 Russia
[pem_pos=2, final_scores=0.95115656,
final_local_scores=0.91874546, pem=-1.0515022,
similarity=37.225143]
4: 25391 Russia
[pem_pos=0, final_scores=0.8650637,
final_local_scores=0.8247021, pem=-0.20299491,
similarity=29.804386]
5: 18998750 Reuters
[pem_pos=0, final_scores=0.33649176,
final_local_scores=0.42480493, pem=0.0, similarity=16.766739]
6: 25391 Russia
[pem_pos=2, final_scores=0.87084454,
final_local_scores=0.8310168, pem=-1.0515022,
similarity=34.607143]
7: 25391 Russia

```
[pem_pos=2, final_scores=0.8826043,
final_local_scores=0.84386253, pem=-1.0515022,
similarity=34.990486]
8: 25391 Russia
    [pem_pos=2, final_scores=0.90649146,
final_local_scores=0.8699556, pem=-1.0515022,
similarity=35.769157]
9: 25391 Russia
    [pem_pos=1, final_scores=0.83623105,
final_local_scores=0.7932068, pem=-1.0515022,
similarity=33.478817]
10: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.73883516,
final_local_scores=0.79772705, pem=0.0, similarity=27.89548]
11: 25391 Russia
    [pem_pos=2, final_scores=0.8203301,
final_local_scores=0.7758373, pem=-1.0515022,
similarity=32.960476]
12: 26779 Soviet_Union
    [pem_pos=0, final_scores=0.40629613,
final_local_scores=0.37310648, pem=-0.075323656,
similarity=15.633575]
13: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.5622814,
final_local_scores=0.60486895, pem=0.0, similarity=22.14021]
14: 25391 Russia
    [pem_pos=2, final_scores=0.8754124,
final_local_scores=0.8360064, pem=-1.0515022,
similarity=34.756042]
15: 25391 Russia
    [pem_pos=2, final_scores=0.7929967,
final_local_scores=0.7459798, pem=-1.0515022,
similarity=32.06947]
16: 25391 Russia
    [pem_pos=2, final_scores=0.81152135,
final_local_scores=0.7662152, pem=-1.0515022,
similarity=32.673336]
17: 25391 Russia
    [pem_pos=2, final_scores=0.8809599,
final_local_scores=0.8420663, pem=-1.0515022,
similarity=34.936882]
```

Sample: b'129_Viacom'

```
[1 1 Viacom 1] plans " Mission " sequel - report . [2 2 LOS A
NGELES 2] 1996-08-22 [3 3 Paramount Pictures 3] is going
ahead with a sequel to the [4 4 Tom Cruise 4] blockbuster , "
Mission : Impossible " and hopes to release it in the summer
of 1998 , [5 5 Daily Variety 5] reported in its Friday
edition . The big-screen version of the spy TV series has
grossed $ 175 million domestically since opening May 22 , and
$ 338 million overseas so far . It 's the biggest success for
Viacom Inc-owned [6 6 Paramount 6] since 1994 's " [7 7 Forre
```

st Gump 7] " . However , many critics complained its plot was incomprehensible . [8 8 Cruise 8] will reprise his roles as star and co-producer , and will soon meet [9 9 Academy Award-winning 9] screenwriter [10 10 William Goldman 10] , who will write the script , the report said . It said " Mission : Impossible " director [11 11 Brian De Palma 11] would have first crack at the sequel , though no deals have been made yet . [12 12 Goldman 12] , whose [13 13 Oscars 13] were for " Butch Cassidy and the Sundance Kid " and " All the President 's Men " , earlier this summer criticised some of the season 's blockbusters . However , he singled out " Mission : Impossible " as an especially entertaining movie , [14 14 Daily Variety 14] said .

```
1: 24580262 Viacom_(original)
2: 18110 Los_Angeles
3: 22918 Paramount_Pictures
4: 31460 Tom_Cruise
5: 22101487 Variety_(magazine)
6: 22918 Paramount_Pictures
7: 41528 Forrest_Gump
8: 31460 Tom_Cruise
9: 324 Academy_Awards
10: 33904 William_Goldman
11: 4216 Brian_De_Palma
12: 33904 William_Goldman
13: 324 Academy_Awards
14: 22101487 Variety_(magazine)
1: 24580262 Viacom_(original)
    [pem_pos=1, final_scores=0.06672113,
final_local_scores=0.1248261, pem=-1.1099129,
similarity=13.850652]
2: 18110 Los_Angeles
    [pem_pos=0, final_scores=0.31803662,
final_local_scores=0.3209248, pem=-0.100652285,
similarity=14.214111]
3: 22918 Paramount_Pictures
    [pem_pos=0, final_scores=0.385786,
final_local_scores=0.33745176, pem=0.0, similarity=14.159946]
4: 31460 Tom_Cruise
    [pem_pos=0, final_scores=0.5386071,
final_local_scores=0.5163755, pem=0.0, similarity=19.499388]
5: 22101487 Variety_(magazine)
    [pem_pos=0, final_scores=0.50016886,
final_local_scores=0.49556124, pem=0.0, similarity=18.878248]
6: 22918 Paramount_Pictures
    [pem_pos=0, final_scores=0.42645913,
final_local_scores=0.381881, pem=-0.357948,
similarity=17.43238]
7: 41528 Forrest_Gump
    [pem_pos=0, final_scores=0.44887418,
final_local_scores=0.4368925, pem=0.0, similarity=17.127457]
8: 31460 Tom_Cruise
```

```
[pem_pos=0, final_scores=0.69932467,
final_local_scores=0.70019615, pem=0.0, similarity=24.984962]
9: 324 Academy_Awards
    [pem_pos=0, final_scores=0.38496178,
final_local_scores=0.34732392, pem=0.0, similarity=14.454552]
10: 33904 William_Goldman
    [pem_pos=0, final_scores=0.53554606,
final_local_scores=0.5356419, pem=0.0, similarity=20.074335]
11: 4216 Brian_De_Palma
    [pem_pos=0, final_scores=0.48400915,
final_local_scores=0.46798974, pem=0.0, similarity=18.05546]
12: 33904 William_Goldman
    [pem_pos=0, final_scores=0.8927948,
final_local_scores=0.9369969, pem=0.0, similarity=32.051567]
13: 324 Academy_Awards
    [pem_pos=0, final_scores=0.5244077,
final_local_scores=0.49964744, pem=0.0, similarity=19.000189]
14: 22101487 Variety_(magazine)
    [pem_pos=0, final_scores=0.5057511,
final_local_scores=0.501659, pem=0.0, similarity=19.060217]
```

Sample: b'130_CRICKET'

CRICKET - [1 1 CRAWLEY 1] FORCED TO SIT AND WAIT . [2 2 LOND
ON 2] 1996-08-23 [3 3 England 3] batsman [4 4 John Crawley 4]
was forced to endure a frustrating delay of over three hours
before resuming his quest for a maiden test century in the
third test against [5 5 Pakistan 5] on Friday . Heavy
overnight rain and morning drizzle ruled out any play before
lunch on the second day but an improvement in the weather
prompted the umpires to announce a 1415 local time (1315 [6 6
GMT 6]) start in the event of no further rain . [7 7 Crawley
7] , unbeaten on 94 overnight in an [8 8 England 8] total of
278 for six , was spotted strumming a guitar in the dressing-
room as the [9 9 Oval 9] ground staff took centre stage .
There were several damp patches on the square and the
outfield and it was still raining when the players took an
early lunch at 1230 local time (1130 [10 10 GMT 10]) . When
brighter weather finally arrived , the umpires announced a
revised figure of 67 overs to be bowled with play extended to
at least 1900 local time (1800 [11 11 GMT 11]) .

```
1: 156618 Crawley
2: 17867 London
3: 759125 England_cricket_team
4: 2930949 John_Crawley
5: 23235 Pakistan
6: 12701 Greenwich_Mean_Time
7: 2930949 John_Crawley
8: 759125 England_cricket_team
9: 387687 The_Oval
10: 12701 Greenwich_Mean_Time
11: 12701 Greenwich_Mean_Time
1: 156618 Crawley
```

```
[pem_pos=1, final_scores=0.3234524,
final_local_scores=0.34806535, pem=-0.047762234,
similarity=14.736416]
2: 17867 London
    [pem_pos=0, final_scores=0.6254132,
final_local_scores=0.67256445, pem=-0.09086737,
similarity=24.654528]
3: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.6739315,
final_local_scores=0.6722074, pem=-3.3054674,
similarity=42.12537]
4: 2930949 John_Crawley
    [pem_pos=0, final_scores=0.70820504,
final_local_scores=0.75407636, pem=0.0, similarity=26.592855]
5: 23235 Pakistan
    [pem_pos=0, final_scores=0.36306137,
final_local_scores=0.41960233, pem=-0.21944217,
similarity=17.804844]
6: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.6815569,
final_local_scores=0.74985266, pem=0.0, similarity=26.466812]
7: 2930949 John_Crawley
    [pem_pos=0, final_scores=0.68185705,
final_local_scores=0.7252951, pem=0.0, similarity=25.733967]
8: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.74907476,
final_local_scores=0.75429004, pem=-3.3054674,
similarity=44.57488]
9: 387687 The_Oval
    [pem_pos=3, final_scores=0.50352985,
final_local_scores=0.49853677, pem=-2.1624494,
similarity=30.726782]
10: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.6889051,
final_local_scores=0.7578795, pem=0.0, similarity=26.706348]
11: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.677414,
final_local_scores=0.7453272, pem=0.0, similarity=26.331762]
```

Sample: b'131_MOTOR'

MOTOR RACING - [1 1 BELGIAN 1] 1 GRAND PRIX PRACTICE TIMES .
[2 2 SPA-FRANCORCHAMPS 2] , [3 3 Belgium 3] 1996-08-23
Leading times after Friday 's opening practice sessions for
Sunday 's [4 4 Belgian 4] Grand Prix motor race : 1. [5 5 Ger
hard Berger 5] ([6 6 Austria 6]) [7 7 Benetton 7] 1 minute
53.706 seconds 2. [8 8 David Coulthard 8] ([9 9 Britain 9])
[10 10 McLaren 10] 1:54.342 3. [11 11 Jacques Villeneuve 11]
([12 12 Canada 12]) [13 13 Williams 13] 1:54.443 4. [14 14 M
ika Hakkinen 14] ([15 15 Finland 15]) [16 16 McLaren 16]
1:54.754 5. [17 17 Heinz-Harald Frentzen 17] ([18 18 Germany
18]) 1:54.984 6. [19 19 Jean Alesi 19] ([20 20 France 20])
[21 21 Benetton 21] 1:55.101 7. [22 22 Damon Hill 22] ([23 23]

Britain 23]) [24 24 Williams 24] 1:55.281 8. [25 25 Michael S
chumacher 25] ([26 26 Germany 26]) 1:55.333 9. [27 27 Martin
Brundle 27] ([28 28 Britain 28]) [29 29 Jordan 29] 1:55.385
10. [30 30 Rubens Barrichello 30] ([31 31 Brazil 31]) [32 32
Jordan 32] 1:55.645 11. [33 33 Johnny Herbert 33] ([34 34 Br
itain 34]) [35 35 Sauber 35] 1:56.318 12. [36 36 Olivier Pan
is 36] ([37 37 France 37]) [38 38 Ligier 38] 1:56.417
1: 3343 Belgium
2: 948477 Circuit_de_Spa-Francorchamps
3: 3343 Belgium
4: 3343 Belgium
5: 343833 Gerhard_Berger
6: 26964606 Austria
7: 445741 Benetton_Formula
8: 19636733 David_Coulthard
9: 31717 United_Kingdom
10: 20994 McLaren
11: 64258 Jacques_Villeneuve
12: 5042916 Canada
13: 34104 Williams_F1
14: 63073 Mika_Hakkinen
15: 10577 Finland
16: 20994 McLaren
17: 756775 Heinz-Harald_Frentzen
18: 11867 Germany
19: 341553 Jean_Alesi
20: 5843419 France
21: 445741 Benetton_Formula
22: 78249 Damon_Hill
23: 31717 United_Kingdom
24: 34104 Williams_F1
25: 20396 Michael_Schumacher
26: 11867 Germany
27: 341560 Martin_Brundle
28: 31717 United_Kingdom
29: 458472 Jordan_Grand_Prix
30: 252989 Rubens_Barrichello
31: 3383 Brazil
32: 458472 Jordan_Grand_Prix
33: 342991 Johnny_Herbert
34: 31717 United_Kingdom
35: 3158734 Sauber
36: 414513 Olivier_Panis
37: 5843419 France
38: 24726570 Equipe_Ligier
1: 10854 Formula_One
[winner: 348674 Grand_Prix_motor_racing, pem_pos=0,
final_scores=-0.13098091, final_local_scores=-0.18103679,
pem=-1.5726227, similarity=7.239381]
gt: 10854 Formula_One, pem_pos=5,
final_scores=-0.14911444, final_local_scores=-0.25018626,
pem=-2.9734752, similarity=12.793876]

```
1: 3343 Belgium
   [pem_pos=0, final_scores=0.3544107,
final_local_scores=0.3661149, pem=-0.12403572,
similarity=15.6898365]
2: 948477 Circuit_de_Spa-Francorchamps
   [pem_pos=0, final_scores=0.37655622,
final_local_scores=0.36129132, pem=0.0, similarity=14.8713665]
3: 3343 Belgium
   [pem_pos=0, final_scores=0.5856957,
final_local_scores=0.6187586, pem=-0.24180195,
similarity=23.869661]
4: 3343 Belgium
   [pem_pos=0, final_scores=0.2353035,
final_local_scores=0.23600844, pem=-0.12403572,
similarity=11.807201]
5: 343833 Gerhard_Berger
   [pem_pos=0, final_scores=0.4946788,
final_local_scores=0.49550724, pem=0.0, similarity=18.876637]
6: 26964606 Austria
   [pem_pos=0, final_scores=0.6235806,
final_local_scores=0.68059844, pem=-0.25262418,
similarity=25.773937]
7: 445741 Benetton_Formula
   [pem_pos=0, final_scores=0.3391282,
final_local_scores=0.3181196, pem=-0.069357015,
similarity=13.96021]
8: 19636733 David_Coulthard
   [pem_pos=0, final_scores=0.6919777,
final_local_scores=0.69716585, pem=0.0, similarity=24.894533]
9: 31717 United_Kingdom
   [pem_pos=0, final_scores=0.5511187,
final_local_scores=0.59867185, pem=-0.61114657,
similarity=25.278784]
10: 20994 McLaren
    [pem_pos=0, final_scores=0.53566307,
final_local_scores=0.51525784, pem=-0.11262372,
similarity=20.0785]
11: 64258 Jacques_Villeneuve
    [pem_pos=0, final_scores=0.4215843,
final_local_scores=0.40126526, pem=0.0, similarity=16.064268]
12: 5042916 Canada
    [pem_pos=0, final_scores=0.7365773,
final_local_scores=0.81826633, pem=-0.11662793,
similarity=29.142653]
13: 34104 Williams_F1
    [pem_pos=0, final_scores=0.6506191,
final_local_scores=0.64907855, pem=-1.0414248,
similarity=29.122942]
14: 63073 Mika_Häkkinen
    [pem_pos=0, final_scores=0.8080713,
final_local_scores=0.82707375, pem=0.0, similarity=28.771244]
15: 10577 Finland
```

```
[pem_pos=0, final_scores=0.74949795,
final_local_scores=0.8306922, pem=-0.2445894,
similarity=30.20934]
16: 20994 McLaren
    [pem_pos=0, final_scores=0.54069656,
final_local_scores=0.5207561, pem=-0.11262372,
similarity=20.24258]
17: 756775 Heinz-Harald_Frentzen
    [pem_pos=0, final_scores=0.87405545,
final_local_scores=0.9068087, pem=0.0, similarity=31.150694]
18: 11867 Germany
    [pem_pos=0, final_scores=0.71541554,
final_local_scores=0.78084815, pem=-0.15529114,
similarity=28.236277]
19: 341553 Jean_Alesi
    [pem_pos=0, final_scores=0.5177648,
final_local_scores=0.51222736, pem=0.0, similarity=19.375599]
20: 5843419 France
    [pem_pos=0, final_scores=0.70801514,
final_local_scores=0.76169145, pem=-0.16240834,
similarity=27.703306]
21: 445741 Benetton_Formula
    [pem_pos=0, final_scores=0.37154275,
final_local_scores=0.35352755, pem=-0.069357015,
similarity=15.016854]
22: 78249 Damon_Hill
    [pem_pos=0, final_scores=0.62715226,
final_local_scores=0.62430537, pem=0.0, similarity=22.72023]
23: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5296343,
final_local_scores=0.5752035, pem=-0.61114657,
similarity=24.578444]
24: 34104 Williams_F1
    [pem_pos=0, final_scores=0.643578,
final_local_scores=0.64138716, pem=-1.0414248,
similarity=28.893415]
25: 20396 Michael_Schumacher
    [pem_pos=0, final_scores=0.6173695,
final_local_scores=0.60824233, pem=0.0, similarity=22.240877]
26: 11867 Germany
    [pem_pos=0, final_scores=0.74811894,
final_local_scores=0.81657153, pem=-0.15529114,
similarity=29.302334]
27: 341560 Martin_Brundle
    [pem_pos=0, final_scores=0.63311106,
final_local_scores=0.6543572, pem=0.0, similarity=23.617037]
28: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.58091027,
final_local_scores=0.6312147, pem=-0.61114657,
similarity=26.249928]
29: 458472 Jordan_Grand_Prix
    [pem_pos=1, final_scores=0.6713032,
```

```
final_local_scores=0.6756958, pem=-2.8876953,
similarity=39.957558]
30: 252989 Rubens_Barrichello
    [pem_pos=0, final_scores=0.49004894,
final_local_scores=0.48078734, pem=0.0, similarity=18.437366]
31: 3383 Brazil
    [pem_pos=0, final_scores=0.6140186,
final_local_scores=0.7020947, pem=-0.21313342,
similarity=26.20067]
32: 458472 Jordan_Grand_Prix
    [pem_pos=1, final_scores=0.62033945,
final_local_scores=0.62002563, pem=-2.8876953,
similarity=38.29625]
33: 342991 Johnny_Herbert
    [pem_pos=0, final_scores=0.72970957,
final_local_scores=0.7497686, pem=0.0, similarity=26.464304]
34: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.59119064,
final_local_scores=0.64244443, pem=-0.61114657,
similarity=26.585047]
35: 3158734 Sauber
    [pem_pos=0, final_scores=0.37287003,
final_local_scores=0.3565616, pem=0.0, similarity=14.730223]
36: 414513 Olivier_Panis
    [pem_pos=0, final_scores=0.55430037,
final_local_scores=0.57236147, pem=0.0, similarity=21.17012]
37: 5843419 France
    [pem_pos=0, final_scores=0.6842178,
final_local_scores=0.7356965, pem=-0.16240834,
similarity=26.927565]
38: 24726570 Equipe_Ligier
    [pem_pos=1, final_scores=0.34248763,
final_local_scores=0.33166778, pem=-0.5983149,
similarity=17.241072]
```

Sample: b'132_TENNIS'

```
TENNIS - RESULTS AT [1 1 TOSHIBA CLASSIC 1] . [ CORRECTED
05:30 [2 2 GMT 2] ] [3 3 CARLSBAD 3] , [4 4 California 4]
1996-08-22 Results from the $ 450,000 [5 5 Toshiba Classic 5]
tennis tournament on Thursday ( prefix number denotes
seeding ) : Quarter-finals 2 - [6 6 Conchita Martinez 6] ( [7
7 Spain 7] ) beat [8 8 Nathalie Tauziat 8] ( [9 9 France 9] )
6-3 6-4 Second round 5 - [10 10 Gabriela Sabatini 10] ( [11 11
Argentina 11] ) beat [12 12 Asa Carlsson 12] ( [13 13 Sweden 1
3] ) 6-1 7-5 [14 14 Katarina Studenikova 14] ( [15 15 Slovakia
15] ) beat 6 - [16 16 Karina Habsudova 16] ( [17 17 Slovakia 1
7] ) 7-6 ( 7-4 ) 6-2 ( Corrects that Habsudova is sixth
seed ) .
```

```
1: 31726664 Southern_California_Open
2: 12701 Greenwich_Mean_Time
3: 108030 Carlsbad,_California
4: 5407 California
```

```
5: 31726664 Southern_California_Open
6: 784152 Conchita_Martínez
7: 26667 Spain
8: 1028366 Nathalie_Tauziat
9: 5843419 France
10: 782187 Gabriela_Sabatini
11: 18951905 Argentina
12: 4038174 Åsa_Svensson
13: 5058739 Sweden
14: 15290332 Katarína_Studeníková
15: 26830 Slovakia
16: 3062115 Karina_Habšudová
17: 26830 Slovakia
1: 31726664 Southern_California_Open
    [pem_pos=0, final_scores=0.40538245,
final_local_scores=0.44178826, pem=0.0, similarity=17.273556]
2: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.5897333,
final_local_scores=0.7242443, pem=0.0, similarity=25.702608]
3: 108030 Carlsbad,_California
    [pem_pos=0, final_scores=0.50167483,
final_local_scores=0.5669391, pem=-0.36353475,
similarity=22.985266]
4: 5407 California
    [pem_pos=0, final_scores=0.60087806,
final_local_scores=0.6896974, pem=-0.08662026,
similarity=25.142714]
5: 31726664 Southern_California_Open
    [pem_pos=0, final_scores=0.53934413,
final_local_scores=0.5881211, pem=0.0, similarity=21.64042]
6: 784152 Conchita_Martínez
    [pem_pos=0, final_scores=0.8625924,
final_local_scores=0.96256095, pem=0.0, similarity=32.814453]
7: 26667 Spain
    [pem_pos=0, final_scores=0.5934153,
final_local_scores=0.6286493, pem=-0.18857051,
similarity=23.875338]
8: 1028366 Nathalie_Tauziat
    [pem_pos=0, final_scores=0.7491595,
final_local_scores=0.86072206, pem=0.0, similarity=29.775375]
9: 5843419 France
    [pem_pos=0, final_scores=0.7180023,
final_local_scores=0.7676054, pem=-0.16240834,
similarity=27.879791]
10: 782187 Gabriela_Sabatini
    [pem_pos=0, final_scores=0.83584946,
final_local_scores=0.9128755, pem=0.0, similarity=31.331738]
11: 18951905 Argentina
    [pem_pos=0, final_scores=0.5921568,
final_local_scores=0.62580365, pem=-0.30003807,
similarity=24.396595]
12: 4038174 Åsa_Svensson
```

```
[pem_pos=0, final_scores=0.842861,
final_local_scores=0.9841427, pem=0.0, similarity=33.458496]
13: 5058739 Sweden
    [pem_pos=0, final_scores=0.75953394,
final_local_scores=0.8121256, pem=-0.2135508,
similarity=29.486483]
14: 15290332 Katarína_Studeníková
    [pem_pos=0, final_scores=0.6483306,
final_local_scores=0.79851437, pem=0.0, similarity=27.918974]
15: 26830 Slovakia
    [pem_pos=0, final_scores=0.6343512,
final_local_scores=0.6587024, pem=-0.36602598,
similarity=25.737213]
16: 3062115 Karina_Habšudová
    [pem_pos=0, final_scores=0.70145315,
final_local_scores=0.8407734, pem=0.0, similarity=29.180069]
17: 26830 Slovakia
    [pem_pos=0, final_scores=0.6178332,
final_local_scores=0.640659, pem=-0.36602598,
similarity=25.198761]
```

```
Sample: b'133_SOCCER'
SOCCER - [1 1 ENGLISH 1] FIRST DIVISION RESULTS . [2 2 LONDON
2] 1996-08-23 Results of [3 3 English 3] first division
soccer matches on Friday : [4 4 Portsmouth 4] 1 [5 5 Queens Pa
rk Rangers 5] 2 [6 6 Tranmere 6] 3 [7 7 Grimsby 7] 2
1: 9316 England
2: 17867 London
3: 9316 England
4: 564301 Portsmouth_F.C.
5: 84588 Queens_Park_Rangers_F.C.
6: 69219 Tranmere_Rovers_F.C.
7: 451129 Grimsby_Town_F.C.
1: 9316 England
    [pem_pos=1, final_scores=0.5607895,
final_local_scores=0.5649349, pem=-1.3527626,
similarity=28.30503]
2: 17867 London
    [pem_pos=0, final_scores=0.6651285,
final_local_scores=0.7038059, pem=-0.09086737,
similarity=25.586836]
3: 9316 England
    [pem_pos=1, final_scores=0.6675901,
final_local_scores=0.68159825, pem=-1.3527626,
similarity=31.786495]
4: 564301 Portsmouth_F.C.
    [pem_pos=1, final_scores=0.6771931,
final_local_scores=0.66523236, pem=-1.2302613,
similarity=30.631924]
5: 84588 Queens_Park_Rangers_F.C.
    [pem_pos=0, final_scores=0.46540993,
final_local_scores=0.43149847, pem=0.0, similarity=16.966488]
```

```
6: 69219 Tranmere_Rovers_F.C.  
    [pem_pos=0, final_scores=0.43948227,  
final_local_scores=0.4117629, pem=-0.6002322,  
similarity=19.641695]  
7: 451129 Grimsby_Town_F.C.  
    [pem_pos=2, final_scores=0.83208436,  
final_local_scores=0.83787864, pem=-2.6069987,  
similarity=43.270947]
```

Sample: b'134_SOCCER'

```
SOCCER - [1 1 SCOTTISH 1] THIRD DIVISION RESULT . [2 2 GLASG  
OW 2] 1996-08-23 Result of a [3 3 Scottish 3] third division  
soccer match on Friday : [4 4 East Stirling 4] 0 Albion 1  
1: 26994 Scotland  
2: 68736 Glasgow  
3: 26994 Scotland  
4: 817230 East_Stirlingshire_F.C.  
1: 26994 Scotland  
    [pem_pos=0, final_scores=0.5404733,  
final_local_scores=0.491615, pem=-0.47951394,  
similarity=21.368156]  
2: 68736 Glasgow  
    [pem_pos=0, final_scores=0.5226714,  
final_local_scores=0.4941848, pem=-0.15604196,  
similarity=19.685753]  
3: 26994 Scotland  
    [pem_pos=0, final_scores=0.5857505,  
final_local_scores=0.5410735, pem=-0.47951394,  
similarity=22.844097]  
4: 817230 East_Stirlingshire_F.C.  
    [pem_pos=0, final_scores=0.57331806,  
final_local_scores=0.6359886, pem=0.0, similarity=23.068882]
```

Sample: b'135_CRICKET'

```
CRICKET - [1 1 ENGLISH COUNTY CHAMPIONSHIP 1] SCORES . [2 2 L  
ONDON 2] 1996-08-23 Close of play scores in four-day [3 3 Eng  
lish County Championship 3] cricket matches on Friday : Third  
day At [4 4 Weston-super-Mare 4] : [5 5 Durham 5] 326 ( D.  
Cox 95 not out , S. Campbell 69 ; G. Rose 7-73 ) . [6 6 Somer  
set 6] 298-6 ( M. Lathwell 85 , R. Harden 65 ) . Second day  
At [7 7 Colchester 7] : [8 8 Gloucestershire 8] 280 ( [9 9 J.  
Russell 9] 63 , A. Symonds 52 ; A. Cowan 5-68 ) . [10 10 Essex  
10] 194-0 ( G. Gooch 105 not out , D. Robinson 72 not out ) .  
At [11 11 Cardiff 11] : [12 12 Kent 12] 255-3 ( D. Fulton  
64 , M. Walker 59 , C. Hooper 52 not out ) v [13 13 Glamorgan  
13] . At [14 14 Leicester 14] : [15 15 Leicestershire 15]  
343-8 ( P. Simmons 108 , P. Nixon 67 not out ) v [16 16 Hamps  
hire 16] . At [17 17 Northampton 17] : [18 18 Sussex 18] 389  
( N. Lenham 145 , V. Drakes 59 , A. Wells 51 ; A. Penberthy  
4-36 ) . [19 19 Northamptonshire 19] 160-4 ( K. Curran 79 not  
out ) . At [20 20 Trent Bridge 20] : [21 21 Nottinghamshire 21  
] 392-6 ( G. Archer 143 not out , M. Dowman 107 ) v [22 22 Su
```

rrey 22] . At [23 23 Worcester 23] : [24 24 Warwickshire 24] 310 (A. Giles 83 , T. Munton 54 not out , W. Khan 52 ; R. Illingworth 4-54 , S. Lampitt 4-90) . [25 25 Worcestershire 25] 10-0 . At [1 1 Headingley 1] : [26 26 Yorkshire 26] 529-8 declared (C. White 181 , R. Blakey 109 not out , M. Moxon 66 , M. Vaughan 57) . [27 27 Lancashire 27] 162-4 (N. Fairbrother 53 not out) .

1: 683599 County_Championship
2: 17867 London
3: 683599 County_Championship
4: 330072 Weston-super-Mare
5: 1176371 Durham_County_Cricket_Club
6: 1622178 Somerset_County_Cricket_Club
7: 264105 Colchester
8: 1622322 Gloucestershire_County_Cricket_Club
9: 18417732 J._Russell_(Essex_cricketer)
10: 1622252 Essex_County_Cricket_Club
11: 5882 Cardiff
12: 1177719 Kent_County_Cricket_Club
13: 1622320 Glamorgan_County_Cricket_Club
14: 51155 Leicester
15: 1622318 Leicestershire_County_Cricket_Club
16: 1919370 Hampshire_County_Cricket_Club
17: 63734 Northampton
18: 1763691 Sussex_County_Cricket_Club
19: 530308 Northamptonshire_County_Cricket_Club
20: 849500 Trent_Bridge
21: 667341 Nottinghamshire_County_Cricket_Club
22: 766575 Surrey_County_Cricket_Club
23: 58681 Worcester
24: 667292 Warwickshire_County_Cricket_Club
25: 1156428 Worcestershire_County_Cricket_Club
26: 869945 Yorkshire_County_Cricket_Club
27: 749919 Lancashire_County_Cricket_Club

1: 437832 Headingley
[winner: 1876060 Headingley_Stadium, pem_pos=0,
final_scores=0.19866301, final_local_scores=0.17331299,
pem=-0.30334172, similarity=10.911338]
gt: 437832 Headingley, pem_pos=1,
final_scores=0.04190658, final_local_scores=0.027741686,
pem=-0.5704622, similarity=8.019842]

1: 683599 County_Championship
[pem_pos=0, final_scores=0.37095916,
final_local_scores=0.31809932, pem=0.0, similarity=13.582431]

2: 17867 London
[pem_pos=0, final_scores=0.6856635,
final_local_scores=0.7472977, pem=-0.09086737,
similarity=26.884718]

3: 683599 County_Championship
[pem_pos=0, final_scores=0.24631481,
final_local_scores=0.18194434, pem=0.0, similarity=9.519295]

4: 330072 Weston-super-Mare

```
[pem_pos=0, final_scores=0.23358141,
final_local_scores=0.24027672, pem=0.0, similarity=11.26005]
5: 1176371 Durham_County_Cricket_Club
    [pem_pos=2, final_scores=0.84813017,
final_local_scores=0.84129435, pem=-2.3952909,
similarity=42.22158]
6: 1622178 Somerset_County_Cricket_Club
    [pem_pos=1, final_scores=0.6968536,
final_local_scores=0.65848076, pem=-2.4139037,
similarity=36.867275]
7: 264105 Colchester
    [pem_pos=0, final_scores=0.21536969,
final_local_scores=0.20480022, pem=-0.1350495,
similarity=10.93578]
8: 1622322 Gloucestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.67082316,
final_local_scores=0.6281617, pem=-1.379793,
similarity=30.33884]
9: 18417732 J. Russell_(Essex_cricketer)
    [pem_pos=7, final_scores=0.24127387,
final_local_scores=0.3581527, pem=-1.7917595,
similarity=24.52157]
10: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.8951598,
final_local_scores=0.87773734, pem=-2.3459642,
similarity=43.040863]
11: 5882 Cardiff
    [pem_pos=0, final_scores=0.6383902,
final_local_scores=0.664338, pem=-0.20401932,
similarity=25.024372]
12: 1177719 Kent_County_Cricket_Club
    [pem_pos=1, final_scores=0.7970939,
final_local_scores=0.7678862, pem=-2.5096242,
similarity=40.6527]
13: 1622320 Glamorgan_County_Cricket_Club
    [pem_pos=1, final_scores=0.67648596,
final_local_scores=0.63803273, pem=-0.7903395,
similarity=27.427872]
14: 51155 Leicester
    [pem_pos=0, final_scores=0.60560507,
final_local_scores=0.6044993, pem=-0.2625849,
similarity=23.557154]
15: 1622318 Leicestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.80066985,
final_local_scores=0.77792436, pem=-1.0465667,
similarity=32.995922]
16: 1919370 Hampshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.75172096,
final_local_scores=0.72433656, pem=-1.9232318,
similarity=36.164192]
17: 63734 Northampton
    [pem_pos=0, final_scores=0.54741865,
```

```
final_local_scores=0.54660636, pem=-0.5110059,
similarity=23.180468]
18: 1763691 Sussex_County_Cricket_Club
    [pem_pos=1, final_scores=0.6515157,
final_local_scores=0.620573, pem=-1.5566913,
similarity=31.074379]
19: 530308 Northamptonshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.39985693,
final_local_scores=0.33848852, pem=-1.2069724,
similarity=20.75459]
20: 849500 Trent_Bridge
    [pem_pos=0, final_scores=0.63013536,
final_local_scores=0.61812747, pem=0.0, similarity=22.53587]
21: 667341 Nottinghamshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6485525,
final_local_scores=0.6037524, pem=-1.1067253,
similarity=28.125433]
22: 766575 Surrey_County_Cricket_Club
    [pem_pos=1, final_scores=0.72521275,
final_local_scores=0.6857305, pem=-1.7952642,
similarity=34.316204]
23: 58681 Worcester
    [pem_pos=0, final_scores=0.54887944,
final_local_scores=0.56536996, pem=-0.59997296,
similarity=24.224226]
24: 667292 Warwickshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.7502226,
final_local_scores=0.716915, pem=-1.1093602,
similarity=31.51676]
25: 1156428 Worcestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.7094638,
final_local_scores=0.6749955, pem=-0.9164197,
similarity=29.216557]
26: 869945 Yorkshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.700398,
final_local_scores=0.6599212, pem=-1.7779728,
similarity=33.45197]
27: 749919 Lancashire_County_Cricket_Club
    [pem_pos=1, final_scores=0.55408293,
final_local_scores=0.50187385, pem=-1.7295884,
similarity=28.4724]
1: 1876060 Headingley_Stadium
    [pem_pos=0, final_scores=0.19866301,
final_local_scores=0.17331299, pem=-0.30334172,
similarity=10.911338]
```

Sample: b'136_CRICKET'

CRICKET - [1 1 POLLOCK 1] HOPES FOR RETURN TO [2 2 WARWICKSHIRE 2] . [3 3 LONDON 3] 1996-08-23 [4 4 South African 4] all-rounder [5 5 Shaun Pollock 5] , forced to cut short his first season with [6 6 Warwickshire 6] to have ankle surgery , has told the [7 7 English 7] county he would like to return later

in his career . [8 8 Pollock 8] , who returns home a month early next week , said : " I would like to come back and play county cricket in the future and I do n't think I would like to swap counties . " Explaining his premature departure was unavoidable , [9 9 Pollock 9] said : " I have been carrying the injury for a while and I hope that by having the surgery now I will be able to last out the new season back home . "

1: 897244 Shaun_Pollock
2: 667292 Warwickshire_County_Cricket_Club
3: 17867 London
4: 17416221 South_Africa
5: 897244 Shaun_Pollock
6: 667292 Warwickshire_County_Cricket_Club
7: 9316 England
8: 897244 Shaun_Pollock
9: 897244 Shaun_Pollock
1: 897244 Shaun_Pollock
 [pem_pos=0, final_scores=1.4647306,
final_local_scores=1.5062151, pem=0.0, similarity=49.038174]
2: 667292 Warwickshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.49465948,
final_local_scores=0.5073957, pem=-1.1093602,
similarity=25.264286]
3: 17867 London
 [pem_pos=0, final_scores=0.5986809,
final_local_scores=0.6979771, pem=-0.09086737,
similarity=25.412893]
4: 17416221 South_Africa
 [pem_pos=0, final_scores=0.2716084,
final_local_scores=0.33758563, pem=-0.1389451,
similarity=14.919547]
5: 897244 Shaun_Pollock
 [pem_pos=0, final_scores=1.1766499,
final_local_scores=1.1847489, pem=0.0, similarity=39.44498]
6: 667292 Warwickshire_County_Cricket_Club
 [pem_pos=1, final_scores=0.6419066,
final_local_scores=0.6682408, pem=-1.1093602,
similarity=30.064224]
7: 9316 England
 [pem_pos=1, final_scores=0.4000708,
final_local_scores=0.47268826, pem=-1.3527626,
similarity=25.552206]
8: 897244 Shaun_Pollock
 [pem_pos=0, final_scores=1.3581201,
final_local_scores=1.3897593, pem=0.0, similarity=45.562904]
9: 897244 Shaun_Pollock
 [pem_pos=0, final_scores=1.2107261,
final_local_scores=1.2287538, pem=0.0, similarity=40.75818]

Sample: b'137_CRICKET'

CRICKET - [1 1 ENGLAND 1] V [2 2 PAKISTAN 2] FINAL TEST
SCOREBOARD . [3 3 LONDON 3] 1996-08-23 Scoreboard on the

second day of the third and final test between [4 4 England 4] and [5 5 Pakistan 5] at [6 6 The Oval 6] on Friday : [7 7 Eng land 7] first innings M. Atherton b [8 8 Waqar Younis 8] 31 A. Stewart b [9 9 Mushtaq Ahmed 9] 44 N. Hussain c [10 10 Sae ed Anwar 10] b [11 11 Waqar Younis 11] 12 G. Thorpe lbw b [12 12 Mohammad Akram 12] 54 J. Crawley b [13 13 Waqar Younis 13] 106 N. Knight b [14 14 Mushtaq Ahmed 14] 17 C. Lewis b [15 15 Wasim Akram 15] 5 I. Salisbury c [16 16 Inzamam-ul-Haq 16] b [17 17 Wasim Akram 17] 5 D. Cork c [18 18 Moin Khan 18] b [19 1 9 Waqar Younis 19] 0 R. Croft not out 5 A. Mullally b [20 20 W asim Akram 20] 24 Extras (lb-12 w-1 nb-10) 23 Total 326 Fall of wickets : 1-64 2-85 3-116 4-205 5-248 6-273 7-283 8-284 9-295 Bowling : [21 21 Wasim Akram 21] 29.2-9-83-3 , [22 22 Waqar Younis 22] 25-6-95-4 , [23 23 Mohammad Akram 23] 12-1-41-1 , [24 24 Mushtaq Ahmed 24] 27-5-78-2 , Aamir Sohail 6-1-17-0 [25 25 Pakistan 25] first innings [26 26 Saeed Anwar 26] not out 116 Aamir Sohail c [27 27 Cork 27] b [28 28 Croft 28] 46 [29 29 Ijaz Ahmed 29] not out 58 Extras (lb-1 nb-8) 9 Total (for one wicket) 229 Fall of wicket - 1-106 To bat : [30 30 Inzamam-ul-Haq 30] , [31 31 Salim Malik 31] , [32 32 Asif Mujtaba 32] , [33 33 Wasim Akram 33] , [34 34 Moin Kh an 34] , [35 35 Mushtaq Ahmed 35] , [36 36 Waqar Younis 36] , Mohammad Akam Bowling (to date) : 1 Lewis 9-1-49-0 , [37 37 Mullally 37] 9-3-28-0 , [38 38 Croft 38] 17-3-42-1 , [39 39 Co rk 39] 7-1-38-0 , 2 Salisbury 14-0-71-0

- 1: 9316 England
- 2: 23235 Pakistan
- 3: 17867 London
- 4: 759125 England_cricket_team
- 5: 759295 Pakistan_national_cricket_team
- 6: 387687 The_Oval
- 7: 759125 England_cricket_team
- 8: 894433 Waqar_Younis
- 9: 896405 Mushtaq_Ahmed
- 10: 894437 Saeed_Anwar
- 11: 894433 Waqar_Younis
- 12: 1958830 Mohammad_Akram
- 13: 894433 Waqar_Younis
- 14: 896405 Mushtaq_Ahmed
- 15: 501393 Wasim_Akram
- 16: 896403 Inzamam-ul-Haq
- 17: 501393 Wasim_Akram
- 18: 554132 Moin_Khan
- 19: 894433 Waqar_Younis
- 20: 501393 Wasim_Akram
- 21: 501393 Wasim_Akram
- 22: 894433 Waqar_Younis
- 23: 1958830 Mohammad_Akram
- 24: 896405 Mushtaq_Ahmed
- 25: 759295 Pakistan_national_cricket_team
- 26: 894437 Saeed_Anwar
- 27: 2036840 Dominic_Cork

```
28: 1160680 Robert_Croft
29: 938863 Ijaz_Ahmed_(cricketer)
30: 896403 Inzamam-ul-Haq
31: 938861 Saleem_Malik
32: 3976606 Asif_Mujtaba
33: 501393 Wasim_Akram
34: 554132 Moin_Khan
35: 896405 Mushtaq_Ahmed
36: 894433 Waqar_Younis
37: 2222543 Alan_Mullally
38: 1160680 Robert_Croft
39: 2036840 Dominic_Cork
1: 8319566 Lewis_Gunn
    [winner: 22743433 Charles_Lewis_(sculler), pem_pos=3,
final_scores=-0.48888558, final_local_scores=-0.38342336,
pem=-1.0986123, similarity=-1.3779805]
    RECALL MISS: 8319566 Lewis_Gunn]
2: 2416655 Ian_Salisbury
    [winner: 72119 Salisbury, pem_pos=0,
final_scores=-0.53736126, final_local_scores=-0.51710427,
pem=-0.80015767, similarity=-6.9903283]
    RECALL MISS: 2416655 Ian_Salisbury]
1: 9316 England
    [pem_pos=0, final_scores=0.35445964,
final_local_scores=0.451864, pem=-0.19053856,
similarity=18.610415]
2: 23235 Pakistan
    [pem_pos=0, final_scores=0.2930892,
final_local_scores=0.36836934, pem=-0.21944217,
similarity=16.27595]
3: 17867 London
    [pem_pos=0, final_scores=0.56492645,
final_local_scores=0.69248587, pem=-0.09086737,
similarity=25.249023]
4: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.78664094,
final_local_scores=0.7791, pem=-3.3054674,
similarity=45.315258]
5: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.84752613,
final_local_scores=0.8384821, pem=-2.7556365,
similarity=44.09727]
6: 387687 The_Oval
    [pem_pos=0, final_scores=0.6165082,
final_local_scores=0.62547266, pem=0.0, similarity=22.755064]
7: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.7641149,
final_local_scores=0.7544938, pem=-3.3054674,
similarity=44.58096]
8: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.70920974,
final_local_scores=0.68770516, pem=0.0, similarity=24.612207]
```

9: 896405 Mushtaq_Ahmed
[pem_pos=0, final_scores=0.56967944,
final_local_scores=0.55337125, pem=0.0, similarity=20.603415]
10: 894437 Saeed_Anwar
[pem_pos=0, final_scores=0.7838616,
final_local_scores=0.7866715, pem=0.0, similarity=27.56556]
11: 894433 Waqar_Younis
[pem_pos=0, final_scores=0.6600132,
final_local_scores=0.63396543, pem=0.0, similarity=23.008507]
12: 1958830 Mohammad_Akram
[pem_pos=0, final_scores=0.51441723,
final_local_scores=0.52836233, pem=0.0, similarity=19.8571]
13: 894433 Waqar_Younis
[pem_pos=0, final_scores=0.71682197,
final_local_scores=0.69602036, pem=0.0, similarity=24.86035]
14: 896405 Mushtaq_Ahmed
[pem_pos=0, final_scores=0.5859876,
final_local_scores=0.5711854, pem=0.0, similarity=21.135025]
15: 501393 Wasim_Akram
[pem_pos=0, final_scores=0.70935434,
final_local_scores=0.6872006, pem=0.0, similarity=24.597149]
16: 896403 Inzamam-ul-Haq
[pem_pos=0, final_scores=0.9057813,
final_local_scores=0.9075412, pem=0.0, similarity=31.172552]
17: 501393 Wasim_Akram
[pem_pos=0, final_scores=0.7003191,
final_local_scores=0.677331, pem=0.0, similarity=24.30262]
18: 554132 Moin_Khan
[pem_pos=0, final_scores=0.7332144,
final_local_scores=0.76472294, pem=0.0, similarity=26.91057]
19: 894433 Waqar_Younis
[pem_pos=0, final_scores=0.698209,
final_local_scores=0.6756885, pem=0.0, similarity=24.253605]
20: 501393 Wasim_Akram
[pem_pos=0, final_scores=0.69483024,
final_local_scores=0.6713352, pem=0.0, similarity=24.123695]
21: 501393 Wasim_Akram
[pem_pos=0, final_scores=0.8077231,
final_local_scores=0.7946536, pem=0.0, similarity=27.80376]
22: 894433 Waqar_Younis
[pem_pos=0, final_scores=0.762543,
final_local_scores=0.74596363, pem=0.0, similarity=26.350756]
23: 1958830 Mohammad_Akram
[pem_pos=0, final_scores=0.716319,
final_local_scores=0.74890935, pem=0.0, similarity=26.438662]
24: 896405 Mushtaq_Ahmed
[pem_pos=0, final_scores=0.60624415,
final_local_scores=0.59331256, pem=0.0, similarity=21.795343]
25: 759295 Pakistan_national_cricket_team
[pem_pos=1, final_scores=0.81082636,
final_local_scores=0.7983932, pem=-2.7556365,
similarity=42.900936]

```
26: 894437 Saeed_Anwar
    [pem_pos=0, final_scores=0.86765414,
final_local_scores=0.87820214, pem=0.0, similarity=30.297016]
27: 2036840 Dominic_Cork
    [pem_pos=8, final_scores=1.1396475,
final_local_scores=1.1826456, pem=-3.606134,
similarity=58.992935]
28: 1160680 Robert_Croft
    [pem_pos=7, final_scores=0.97817963,
final_local_scores=1.035461, pem=-3.16984,
similarity=52.22802]
29: 938863 Ijaz_Ahmed_(cricketer)
    [pem_pos=0, final_scores=0.61321044,
final_local_scores=0.60634947, pem=0.0, similarity=22.184391]
30: 896403 Inzamam-ul-Haq
    [pem_pos=0, final_scores=1.0564989,
final_local_scores=1.0721774, pem=0.0, similarity=36.08563]
31: 938861 Saleem_Malik
    [pem_pos=0, final_scores=0.7417844,
final_local_scores=0.7658922, pem=0.0, similarity=26.945463]
32: 3976606 Asif_Mujtaba
    [pem_pos=0, final_scores=0.7706476,
final_local_scores=0.80907017, pem=0.0, similarity=28.23398]
33: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.77039355,
final_local_scores=0.7538767, pem=0.0, similarity=26.586897]
34: 554132 Moin_Khan
    [pem_pos=0, final_scores=0.79036313,
final_local_scores=0.8271493, pem=0.0, similarity=28.773497]
35: 896405 Mushtaq_Ahmed
    [pem_pos=0, final_scores=0.64588684,
final_local_scores=0.63661623, pem=0.0, similarity=23.08761]
36: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.7825244,
final_local_scores=0.7677903, pem=0.0, similarity=27.002108]
37: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.6424193,
final_local_scores=0.6739463, pem=-0.73930055,
similarity=28.222048]
38: 1160680 Robert_Croft
    [pem_pos=7, final_scores=0.78475493,
final_local_scores=0.8241738, pem=-3.16984,
similarity=45.922787]
39: 2036840 Dominic_Cork
    [pem_pos=8, final_scores=0.8559275,
final_local_scores=0.8727246, pem=-3.606134,
similarity=49.744274]
```

Sample: b'138_CRICKET'

CRICKET - [1 1 ENGLAND 1] 326 ALL OUT V [2 2 PAKISTAN 2] IN
THIRD TEST . [3 3 LONDON 3] 1996-08-23 [4 4 England 4] were
all out for 326 in their first innings on the second day of

the third and final test against [5 5 Pakistan 5] at [6 6 The Oval 6] on Friday . Score : [7 7 England 7] 326 (J. Crawley 106 , G. Thorpe 54 . [8 8 Waqar Younis 8] 4-95)

1: 759125 England_cricket_team
2: 759295 Pakistan_national_cricket_team
3: 17867 London
4: 759125 England_cricket_team
5: 759295 Pakistan_national_cricket_team
6: 387687 The_Oval
7: 759125 England_cricket_team
8: 894433 Waqar_Younis
1: 759125 England_cricket_team
 [pem_pos=2, final_scores=0.7024035,
final_local_scores=0.65532297, pem=-3.3054674,
similarity=41.621502]
2: 759295 Pakistan_national_cricket_team
 [pem_pos=1, final_scores=0.609458,
final_local_scores=0.5706534, pem=-2.7556365,
similarity=36.10473]
3: 17867 London
 [pem_pos=0, final_scores=0.58425266,
final_local_scores=0.6719662, pem=-0.09086737,
similarity=24.636675]
4: 759125 England_cricket_team
 [pem_pos=2, final_scores=0.763105,
final_local_scores=0.72163016, pem=-3.3054674,
similarity=43.600243]
5: 759295 Pakistan_national_cricket_team
 [pem_pos=1, final_scores=0.87335724,
final_local_scores=0.85892314, pem=-2.7556365,
similarity=44.70727]
6: 387687 The_Oval
 [pem_pos=0, final_scores=0.6423488,
final_local_scores=0.6324652, pem=0.0, similarity=22.963736]
7: 759125 England_cricket_team
 [pem_pos=2, final_scores=0.7364624,
final_local_scores=0.6925271, pem=-3.3054674,
similarity=42.73175]
8: 894433 Waqar_Younis
 [pem_pos=0, final_scores=0.52288306,
final_local_scores=0.54062665, pem=0.0, similarity=20.223091]

Sample: b'139_SOCCER'
SOCCER - SPONSORS CASH IN ON [1 1 RAVANELLI 1] 'S SHIRT
DANCE . [2 2 LONDON 2] 1996-08-23 [3 3 Middlesbrough 3] 's [4
4 Italian 4] striker [5 5 Fabrizio Ravanelli 5] is to wear
his team sponsor 's name on the inside of his shirt so it can
be seen when he scores . Every time he finds the net , the
grey-haired forward pulls his shirtfront over his head as he
runs to salute the fans , and [6 6 Middlesbrough 6] 's
sponsors want to cash in on the spectacle . " Having seen [7 7
Ravanelli 7] celebrate his goals ... we thought it would be

fun to have (the name) on the inside of his shirt , " a spokesman for the sponsors said . " It will give the fans something else to look at besides his chest . " [8 8 Ravanelli 8] aggravated a foot injury in the 1-0 defeat at [9 9 Chelsea 9] on Wednesday and was given only an even chance of playing at [10 10 Nottingham Forest 10] on Saturday by his manager [11 11 Bryan Robson 11] .

1: 2189648 Fabrizio_Ravanelli
2: 17867 London
3: 298603 Middlesbrough_F.C.
4: 14532 Italy
5: 2189648 Fabrizio_Ravanelli
6: 298603 Middlesbrough_F.C.
7: 2189648 Fabrizio_Ravanelli
8: 2189648 Fabrizio_Ravanelli
9: 7473 Chelsea_F.C.
10: 157294 Nottingham_Forest_F.C.
11: 345428 Bryan_Robson
1: 2189648 Fabrizio_Ravanelli
[pem_pos=0, final_scores=0.7696721,
final_local_scores=0.77066225, pem=0.0, similarity=27.087812]
2: 17867 London
[pem_pos=0, final_scores=0.5199982,
final_local_scores=0.6144326, pem=-0.09086737,
similarity=22.919756]
3: 298603 Middlesbrough_F.C.
[pem_pos=0, final_scores=0.57168955,
final_local_scores=0.5456905, pem=-0.5175458,
similarity=23.1887]
4: 14532 Italy
[pem_pos=0, final_scores=0.5071924,
final_local_scores=0.63115263, pem=-0.75948656,
similarity=27.054771]
5: 2189648 Fabrizio_Ravanelli
[pem_pos=0, final_scores=0.70305336,
final_local_scores=0.6978914, pem=0.0, similarity=24.916185]
6: 298603 Middlesbrough_F.C.
[pem_pos=0, final_scores=0.56901985,
final_local_scores=0.5427743, pem=-0.5175458,
similarity=23.101675]
7: 2189648 Fabrizio_Ravanelli
[pem_pos=0, final_scores=0.62564635,
final_local_scores=0.613336, pem=0.0, similarity=22.392883]
8: 2189648 Fabrizio_Ravanelli
[pem_pos=0, final_scores=0.7517658,
final_local_scores=0.7511024, pem=0.0, similarity=26.504107]
9: 7473 Chelsea_F.C.
[pem_pos=0, final_scores=0.72734207,
final_local_scores=0.71261656, pem=-0.51707125,
similarity=28.167526]
10: 157294 Nottingham_Forest_F.C.
[pem_pos=0, final_scores=0.49381953,

```
final_local_scores=0.4601395, pem=0.0, similarity=17.821194]
11: 345428 Bryan Robson
    [pem_pos=0, final_scores=0.5347672,
final_local_scores=0.5438349, pem=0.0, similarity=20.318832]
```

Sample: b'140_TENNIS'

TENNIS - [1 1 AUSTRALIANS 1] ADVANCE AT [2 2 CANADIAN OPEN 2]. [3 3 TORONTO 3] 1996-08-22 It was Australia Day at the \$ 2 million [4 4 Canadian Open 4] on Thursday as three Aussies reached the quarter-finals with straight-set victories . Unseeded [5 5 Patrick Rafter 5] recorded the most noteworthy result as he upset sixth-seeded [6 6 American 6] [7 7 MaliVai Washington 7] 6-2 6-1 in just 50 minutes . [8 8 Todd Woodbridge 8] , who defeated [9 9 Canadian 9] [10 10 Daniel Nestor 10] 7-6 (7-2) 7-6 (7-4) , and [11 11 Mark Philippoussis 11] , a 6-3 6-4 winner over [12 12 Bohdan Ulihrach 12] of the [13 13 Czech Republic 13] , also advanced and will meet in Friday 's quarter-finals . Third-seeded [14 14 Wayne Ferreira 14] of [15 15 South Africa 15] defeated [16 16 Tim Henman 16] of [17 17 Britain 17] 6-4 6-4 after a three-hour evening rain delay and fifth-seeded [18 18 Thomas Enqvist 18] of [19 19 Sweden 19] won his third-round match , eliminating [20 20 Petr Korda 20] of the [21 21 Czech Republic 21] 6-3 6-4 . [22 22 Ferreira 22] and [23 23 Enqvist 23] play in a Friday night quarter-final . Two [24 24 Americans 24] , seventh seed [25 25 Todd Martin 25] and unseeded Alex O'Brien , will meet on Friday after winning matches on Thursday . [26 26 Martin 26] overcame [27 27 Cedric Pioline 27] of [28 28 France 28] 2-6 6-2 6-4 and O'Brien beat [29 29 Mikael Tillstrom 29] of [30 30 Sweden 30] 6-3 2-6 6-3 . " If you really look at the match , " said the 12th-ranked [31 31 Washington 31] after losing to the 70th-ranked [32 32 Rafter 32] , " I never really got a chance to play because he was serving big and getting in close to the net . " He was also able to handle my serve pretty easily because my (first) service percentage was only in the 40s . Put those two things together and you get a loss . " [33 33 Rafter 33] missed 10 weeks after wrist surgery earlier this year and the time away from tennis has given him a new perspective . " Before when I was on tour , I always felt I had to be in bed by 9:30 or 10 o'clock and I had to be up at a certain time , " [34 34 Rafter 34] said . " Now I can go to bed at at midnight or wake up at seven in the morning . I just do n't have as many set routines and it 's made me a happier person . " [35 35 Martin 35] was pleased with his victory over [36 36 Pioline 36] , his first in five meetings with the 11th-ranked Frenchman . " It 's always difficult to win a match when you lose the first set , especially against someone you have never beaten , " he said . " I got more aggressive in the second and third sets and the wind picked up and that also affected things because Cedric definitely went off a little bit . " The 26-year-old O'Brien , who won the ATP Tour stop in [37 37 New Haven 37] last week , has now

won 18 of his last 20 matches , dating back to qualifying rounds in [38 38 Los Angeles 38] in late July . He ranks 76th after being 285th four weeks ago . " I feel I 'm hitting the ball well even though I 'm having more mental letdowns than I did last week , " O'Brien said . " But I 'm still competing well . " " I got a lot of first serves in , " said [39 39 Enqvist 39] about his victory over [40 40 Korda 40] . " I didn't miss that many shots and he was making the mistakes . " Still marvelling at an exciting 64-stroke rally he won in the last game of his second-round match against [41 41 Javier Sanchez 41] of [42 42 Spain 42] on Tuesday , [43 43 Enqvist 43] joked , " Today against [44 44 Petr 44] there were about 64 strokes in the whole match . It was mostly short points . "

- 1: 4689264 Australia
- 2: 1872219 Canadian_Open_(tennis)
- 3: 64646 Toronto
- 4: 1872219 Canadian_Open_(tennis)
- 5: 23297 Patrick_Rafter
- 6: 3434750 United_States
- 7: 873616 MaliVai_Washington
- 8: 1547951 Todd_Woodbridge
- 9: 5042916 Canada
- 10: 998494 Daniel_Nestor
- 11: 261368 Mark_Philippoussis
- 12: 2805841 Bohdan_Ulihrach
- 13: 5321 Czech_Republic
- 14: 1723516 Wayne_Ferreira
- 15: 17416221 South_Africa
- 16: 262030 Tim_Henman
- 17: 31717 United_Kingdom
- 18: 933709 Thomas_Enqvist
- 19: 5058739 Sweden
- 20: 1439769 Petr_Korda
- 21: 5321 Czech_Republic
- 22: 1723516 Wayne_Ferreira
- 23: 933709 Thomas_Enqvist
- 24: 3434750 United_States
- 25: 871725 Todd_Martin
- 26: 871725 Todd_Martin
- 27: 979076 Cédric_Pioline
- 28: 5843419 France
- 29: 3943969 Mikael_Tillström
- 30: 5058739 Sweden
- 31: 873616 MaliVai_Washington
- 32: 23297 Patrick_Rafter
- 33: 23297 Patrick_Rafter
- 34: 23297 Patrick_Rafter
- 35: 871725 Todd_Martin
- 36: 979076 Cédric_Pioline
- 37: 53825 New_Haven,_Connecticut
- 38: 18110 Los_Angeles
- 39: 933709 Thomas_Enqvist

40: 1439769 Petr_Korda
41: 420527 Javier_Sánchez
42: 26667 Spain
43: 933709 Thomas_Enqvist
44: 1439769 Petr_Korda
1: 4689264 Australia
[pem_pos=1, final_scores=0.49499506,
final_local_scores=0.588155, pem=-1.143797,
similarity=27.861576]
2: 1872219 Canadian_Open_(tennis)
[pem_pos=0, final_scores=0.57405436,
final_local_scores=0.5831527, pem=-0.58982956,
similarity=24.699738]
3: 64646 Toronto
[pem_pos=0, final_scores=0.56638986,
final_local_scores=0.65450525, pem=-0.12291032,
similarity=24.289862]
4: 1872219 Canadian_Open_(tennis)
[pem_pos=0, final_scores=0.40942377,
final_local_scores=0.40331876, pem=-0.58982956,
similarity=19.333134]
5: 23297 Patrick_Rafter
[pem_pos=0, final_scores=0.9099559,
final_local_scores=0.9197379, pem=0.0, similarity=31.536526]
6: 3434750 United_States
[pem_pos=0, final_scores=0.6065628,
final_local_scores=0.6978041, pem=-0.12994486,
similarity=25.62024]
7: 873616 MaliVai_Washington
[pem_pos=0, final_scores=0.7307419,
final_local_scores=0.76232105, pem=0.0, similarity=26.838894]
8: 1547951 Todd_Woodbridge
[pem_pos=0, final_scores=0.69152254,
final_local_scores=0.7085862, pem=0.0, similarity=25.235338]
9: 5042916 Canada
[pem_pos=0, final_scores=0.6516894,
final_local_scores=0.75196636, pem=-0.41303828,
similarity=28.776054]
10: 998494 Daniel_Nestor
[pem_pos=0, final_scores=0.9212952,
final_local_scores=0.9605246, pem=0.0, similarity=32.753685]
11: 261368 Mark_Philippoussis
[pem_pos=0, final_scores=0.89296967,
final_local_scores=0.9147815, pem=0.0, similarity=31.388617]
12: 2805841 Bohdan_Ulihrach
[pem_pos=0, final_scores=0.7520744,
final_local_scores=0.8305296, pem=0.0, similarity=28.874372]
13: 5321 Czech_Republic
[pem_pos=0, final_scores=0.48729312,
final_local_scores=0.5538091, pem=-0.27794555,
similarity=22.127993]
14: 1723516 Wayne_Ferreira

[pem_pos=0, final_scores=0.76866466,
final_local_scores=0.800249, pem=0.0, similarity=27.97074]
15: 17416221 South_Africa
[pem_pos=0, final_scores=0.46253532,
final_local_scores=0.57455283, pem=-0.26251382,
similarity=22.663107]
16: 262030 Tim_Henman
[pem_pos=0, final_scores=0.8061951,
final_local_scores=0.807718, pem=0.0, similarity=28.193628]
17: 31717 United_Kingdom
[pem_pos=0, final_scores=0.4939286,
final_local_scores=0.59687597, pem=-0.61114657,
similarity=25.225191]
18: 933709 Thomas_Enqvist
[pem_pos=0, final_scores=0.7739083,
final_local_scores=0.7830987, pem=0.0, similarity=27.45894]
19: 5058739 Sweden
[pem_pos=0, final_scores=0.698721,
final_local_scores=0.7886272, pem=-0.2135508,
similarity=28.785244]
20: 1439769 Petr_Korda
[pem_pos=0, final_scores=0.91561335,
final_local_scores=0.9494783, pem=0.0, similarity=32.42404]
21: 5321 Czech_Republic
[pem_pos=0, final_scores=0.44690114,
final_local_scores=0.509687, pem=-0.27794555,
similarity=20.8113]
22: 1723516 Wayne_Ferreira
[pem_pos=0, final_scores=0.8666981,
final_local_scores=0.91184896, pem=0.0, similarity=31.301104]
23: 933709 Thomas_Enqvist
[pem_pos=0, final_scores=0.7591582,
final_local_scores=0.7669863, pem=0.0, similarity=26.978115]
24: 3434750 United_States
[pem_pos=0, final_scores=0.35021752,
final_local_scores=0.41778582, pem=-0.5581965,
similarity=19.592834]
25: 871725 Todd_Martin
[pem_pos=0, final_scores=0.88372856,
final_local_scores=0.9010431, pem=0.0, similarity=30.978636]
26: 871725 Todd_Martin
[pem_pos=0, final_scores=0.90292495,
final_local_scores=0.92201227, pem=0.0, similarity=31.604397]
27: 979076 Cédric_Pioline
[pem_pos=0, final_scores=0.9007599,
final_local_scores=0.94084543, pem=0.0, similarity=32.166416]
28: 5843419 France
[pem_pos=0, final_scores=0.656701,
final_local_scores=0.74493635, pem=-0.16240834,
similarity=27.2033]
29: 3943969 Mikael_Tillström
[pem_pos=0, final_scores=0.7487808,

```
final_local_scores=0.8234407, pem=0.0, similarity=28.662825]
30: 5058739 Sweden
    [pem_pos=0, final_scores=0.6395051,
final_local_scores=0.7239428, pem=-0.2135508,
similarity=26.854933]
31: 873616 MaliVai_Washington
    [pem_pos=0, final_scores=0.76202995,
final_local_scores=0.79649854, pem=0.0, similarity=27.858818]
32: 23297 Patrick_Rafter
    [pem_pos=0, final_scores=0.80243,
final_local_scores=0.8051739, pem=0.0, similarity=28.117708]
33: 23297 Patrick_Rafter
    [pem_pos=0, final_scores=0.9407074,
final_local_scores=0.9562209, pem=0.0, similarity=32.62525]
34: 23297 Patrick_Rafter
    [pem_pos=0, final_scores=0.86974424,
final_local_scores=0.8787045, pem=0.0, similarity=30.312008]
35: 871725 Todd_Martin
    [pem_pos=0, final_scores=1.0777494,
final_local_scores=1.1129816, pem=0.0, similarity=37.303303]
36: 979076 Cédric_Pioline
    [pem_pos=0, final_scores=0.8705048,
final_local_scores=0.9077963, pem=0.0, similarity=31.180164]
37: 53825 New_Haven,_Connecticut
    [pem_pos=0, final_scores=0.3208378,
final_local_scores=0.40636176, pem=-0.46266162,
similarity=18.732386]
38: 18110 Los_Angeles
    [pem_pos=0, final_scores=0.5009275,
final_local_scores=0.58052427, pem=-0.100652285,
similarity=21.961077]
39: 933709 Thomas_Enqvist
    [pem_pos=0, final_scores=0.6984855,
final_local_scores=0.7007107, pem=0.0, similarity=25.000319]
40: 1439769 Petr_Korda
    [pem_pos=0, final_scores=0.9171296,
final_local_scores=0.95514375, pem=0.0, similarity=32.59311]
41: 420527 Javier_Sánchez
    [pem_pos=0, final_scores=0.677272,
final_local_scores=0.7162032, pem=0.0, similarity=25.462645]
42: 26667 Spain
    [pem_pos=0, final_scores=0.5836015,
final_local_scores=0.67500514, pem=-0.18857051,
similarity=25.258688]
43: 933709 Thomas_Enqvist
    [pem_pos=0, final_scores=0.59755427,
final_local_scores=0.59045863, pem=0.0, similarity=21.710176]
44: 1439769 Petr_Korda
    [pem_pos=0, final_scores=0.59377295,
final_local_scores=0.5979166, pem=0.0, similarity=21.932737]
```

Sample: b'141_TENNIS'

TENNIS - RESULTS AT [1 1 CANADIAN OPEN 1] . [2 2 TORONTO 2]
1996-08-22 Results from the [3 3 Canadian Open 3] tennis
tournament on Thursday (prefix number denotes seeding) :
Third round 3 - [4 4 Wayne Ferreira 4] ([5 5 South Africa 5]
) beat [6 6 Tim Henman 6] ([7 7 Britain 7]) 6-4 6-4 4 - [8 8
Marcelo Rios 8] ([9 9 Chile 9]) beat [10 10 Daniel Vacek 10]
([11 11 Czech Republic 11]) 6-4 6-3 5 - [12 12 Thomas Enqvist 12]
([13 13 Sweden 13]) beat [14 14 Petr Korda 14] ([15 15
Czech Republic 15]) 6-3 6-4 [16 16 Patrick Rafter 16] ([17 17
Australia 17]) beat 6 - [18 18 MaliVai Washington 18] ([19 19
U.S. 19]) 6-2 6-1 7 - [20 20 Todd Martin 20] ([21 21
U.S. 21]) beat 9 - [22 22 Cedric Pioline 22] ([23 23 France 23])
2-6 6-2 6-4 [24 24 Mark Philippoussis 24] ([25 25 Australia 25])
beat [26 26 Bohdan Ulihrach 26] ([27 27 Czech Republic 27]) 6-3 6-4 Alex O'Brien ([28 28 U.S. 28]) beat [29 29
Mikael Tillstrom 29] ([30 30 Sweden 30]) 6-3 2-6 6-3 [31 31
Todd Woodbridge 31] ([32 32 Australia 32]) beat [33 33 Daniel Nestor 33] ([34 34 Canada 34]) 7-6 (7-2) 7-6 (7-4)
1: 1872219 Canadian_Open_(tennis)
2: 64646 Toronto
3: 1872219 Canadian_Open_(tennis)
4: 1723516 Wayne_Ferreira
5: 17416221 South_Africa
6: 262030 Tim_Henman
7: 31717 United_Kingdom
8: 483554 Marcelo_Rios
9: 5489 Chile
10: 3930665 Daniel_Vacek
11: 5321 Czech_Republic
12: 933709 Thomas_Enqvist
13: 5058739 Sweden
14: 1439769 Petr_Korda
15: 5321 Czech_Republic
16: 23297 Patrick_Rafter
17: 4689264 Australia
18: 873616 MaliVai_Washington
19: 3434750 United_States
20: 871725 Todd_Martin
21: 3434750 United_States
22: 979076 Cédric_Pioline
23: 5843419 France
24: 261368 Mark_Philippoussis
25: 4689264 Australia
26: 2805841 Bohdan_Ulihrach
27: 5321 Czech_Republic
28: 3434750 United_States
29: 3943969 Mikael_Tillström
30: 5058739 Sweden
31: 1547951 Todd_Woodbridge
32: 4689264 Australia
33: 998494 Daniel_Nestor
34: 5042916 Canada

```
1: 1872219 Canadian_Open_(tennis)
   [pem_pos=0, final_scores=0.52690226,
final_local_scores=0.55817413, pem=-0.58982956,
similarity=23.954327]
2: 64646 Toronto
   [pem_pos=0, final_scores=0.5739937,
final_local_scores=0.6162472, pem=-0.12291032,
similarity=23.148165]
3: 1872219 Canadian_Open_(tennis)
   [pem_pos=0, final_scores=0.44668096,
final_local_scores=0.47054458, pem=-0.58982956,
similarity=21.339287]
4: 1723516 Wayne_Ferreira
   [pem_pos=0, final_scores=0.81720954,
final_local_scores=0.9088455, pem=0.0, similarity=31.211473]
5: 17416221 South_Africa
   [pem_pos=0, final_scores=0.578159,
final_local_scores=0.64045864, pem=-0.26251382,
similarity=24.629868]
6: 262030 Tim_Henman
   [pem_pos=0, final_scores=0.7490647,
final_local_scores=0.7812193, pem=0.0, similarity=27.402855]
7: 31717 United_Kingdom
   [pem_pos=0, final_scores=0.5168999,
final_local_scores=0.55375093, pem=-0.61114657,
similarity=23.938255]
8: 483554 Marcelo_Ríos
   [pem_pos=0, final_scores=0.67311347,
final_local_scores=0.71306914, pem=0.0, similarity=25.369118]
9: 5489 Chile
   [pem_pos=0, final_scores=0.65386325,
final_local_scores=0.7152166, pem=-0.27404457,
similarity=26.923498]
10: 3930665 Daniel_Vacek
    [pem_pos=0, final_scores=0.7788487,
final_local_scores=0.8846609, pem=0.0, similarity=30.489758]
11: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.5509367,
final_local_scores=0.565815, pem=-0.27794555,
similarity=22.48627]
12: 933709 Thomas_Enqvist
    [pem_pos=0, final_scores=0.8818343,
final_local_scores=0.953481, pem=0.0, similarity=32.543488]
13: 5058739 Sweden
    [pem_pos=0, final_scores=0.7485859,
final_local_scores=0.7874778, pem=-0.2135508,
similarity=28.750942]
14: 1439769 Petr_Korda
    [pem_pos=0, final_scores=0.96186024,
final_local_scores=1.0561726, pem=0.0, similarity=35.608013]
15: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.57478565,
```

```
final_local_scores=0.5918663, pem=-0.27794555,  
similarity=23.263695]  
16: 23297 Patrick_Rafter  
    [pem_pos=0, final_scores=0.9648277,  
final_local_scores=1.0359743, pem=0.0, similarity=35.00525]  
17: 4689264 Australia  
    [pem_pos=0, final_scores=0.6975189,  
final_local_scores=0.74085855, pem=-0.17168477,  
similarity=27.13206]  
18: 873616 MaliVai_Washington  
    [pem_pos=0, final_scores=0.8737069,  
final_local_scores=0.9468507, pem=0.0, similarity=32.345627]  
19: 3434750 United_States  
    [pem_pos=0, final_scores=0.529269,  
final_local_scores=0.5472588, pem=-0.14075686,  
similarity=21.186466]  
20: 871725 Todd_Martin  
    [pem_pos=0, final_scores=0.8731182,  
final_local_scores=0.9378851, pem=0.0, similarity=32.078075]  
21: 3434750 United_States  
    [pem_pos=0, final_scores=0.4867112,  
final_local_scores=0.50077087, pem=-0.14075686,  
similarity=19.799171]  
22: 979076 Cédric_Pioline  
    [pem_pos=0, final_scores=0.80673915,  
final_local_scores=0.8824758, pem=0.0, similarity=30.424551]  
23: 5843419 France  
    [pem_pos=0, final_scores=0.7032337,  
final_local_scores=0.7409051, pem=-0.16240834,  
similarity=27.083]  
24: 261368 Mark_Philippoussis  
    [pem_pos=0, final_scores=0.82309014,  
final_local_scores=0.87669593, pem=0.0, similarity=30.25207]  
25: 4689264 Australia  
    [pem_pos=0, final_scores=0.66004175,  
final_local_scores=0.6999205, pem=-0.17168477,  
similarity=25.910385]  
26: 2805841 Bohdan_Ulihrach  
    [pem_pos=0, final_scores=0.7615532,  
final_local_scores=0.8774031, pem=0.0, similarity=30.273172]  
27: 5321 Czech_Republic  
    [pem_pos=0, final_scores=0.5540995,  
final_local_scores=0.56926984, pem=-0.27794555,  
similarity=22.58937]  
28: 3434750 United_States  
    [pem_pos=0, final_scores=0.48307115,  
final_local_scores=0.49679464, pem=-0.14075686,  
similarity=19.680513]  
29: 3943969 Mikael_Tillström  
    [pem_pos=0, final_scores=0.6662002,  
final_local_scores=0.7613772, pem=0.0, similarity=26.810728]  
30: 5058739 Sweden
```

```
[pem_pos=0, final_scores=0.7590626,
final_local_scores=0.79892206, pem=-0.2135508,
similarity=29.092463]
31: 1547951 Todd_Woodbridge
    [pem_pos=0, final_scores=0.794692,
final_local_scores=0.8728668, pem=0.0, similarity=30.1378]
32: 4689264 Australia
    [pem_pos=0, final_scores=0.6631003,
final_local_scores=0.7032615, pem=-0.17168477,
similarity=26.010088]
33: 998494 Daniel_Nestor
    [pem_pos=0, final_scores=0.8447686,
final_local_scores=0.92651385, pem=0.0, similarity=31.738735]
34: 5042916 Canada
    [pem_pos=0, final_scores=0.6975903,
final_local_scores=0.73488766, pem=-0.11662793,
similarity=26.654467]
```

Sample: b'142_RUGBY'

```
[1 1 RUGBY UNION 1] - [2 2 MULDER 2] OUT OF SECOND TEST . [3
3 JOHANNESBURG 3] 1996-08-23 Centre [4 4 Japie Mulder 4] has
been ruled out of [5 5 South Africa 5] 's team for the second
test against [6 6 New Zealand 6] in [7 7 Pretoria 7] on
Saturday . [8 8 Mulder 8] missed the first test in [9 9 Durban
9] with back spasms and failed a fitness check on Thursday .
But new [10 10 Springbok 10] skipper [11 11 Gary Teichmann 11]
has recovered from a bruised thigh and is ready to play ,
coach [12 12 Andre Markgraaff 12] said . [13 13 Mulder 13] 's
absence means that [14 14 Northern Transvaal 14] centre [15 15
Andre Snyman 15] should win his second cap alongside
provincial colleague Danie van Schalkwyk . Wing Pieter
Hendriks is expected to retain his place , following
speculation that [16 16 Snyman 16] would be picked out of
position on the wing . The line-up would not be announced
until shortly before the start , [17 17 Markgraaff 17] said .
```

```
1: 25405 Rugby_union
2: 21833605 Japie_Mulder
3: 58893 Johannesburg
4: 21833605 Japie_Mulder
5: 17416221 South_Africa
6: 154957 New_Zealand_national_rugby_union_team
7: 25002 Pretoria
8: 21833605 Japie_Mulder
9: 313650 Durban
10: 376115 South_Africa_national_rugby_union_team
11: 2485308 Gary_Teichmann
12: 1709031 Andre_Markgraaff
13: 21833605 Japie_Mulder
14: 8904211 Northerns_cricket_team
15: 8357907 Andre_Snyman
16: 8357907 Andre_Snyman
17: 1709031 Andre_Markgraaff
```

```
1: 25405 Rugby_union
    [pem_pos=0, final_scores=0.45155907,
final_local_scores=0.48855072, pem=0.0, similarity=18.66904]
2: 21833605 Japie_Mulder
    [pem_pos=0, final_scores=1.0752606,
final_local_scores=1.1499343, pem=0.0, similarity=38.406048]
3: 58893 Johannesburg
    [pem_pos=0, final_scores=0.48044574,
final_local_scores=0.5049521, pem=-0.09562059,
similarity=19.67849]
4: 21833605 Japie_Mulder
    [pem_pos=0, final_scores=0.75882846,
final_local_scores=0.7924454, pem=0.0, similarity=27.737865]
5: 17416221 South_Africa
    [pem_pos=0, final_scores=0.4744324,
final_local_scores=0.5213861, pem=-0.26251382,
similarity=21.076506]
6: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=4, final_scores=0.5207048,
final_local_scores=0.5394802, pem=-3.8192244,
similarity=40.958412]
7: 25002 Pretoria
    [pem_pos=0, final_scores=0.5165792,
final_local_scores=0.5374411, pem=0.0, similarity=20.128027]
8: 21833605 Japie_Mulder
    [pem_pos=0, final_scores=1.1682664,
final_local_scores=1.251529, pem=0.0, similarity=41.437836]
9: 313650 Durban
    [pem_pos=0, final_scores=0.51448184,
final_local_scores=0.54589975, pem=-0.07463237,
similarity=20.786312]
10: 376115 South_Africa_national_rugby_union_team
    [pem_pos=1, final_scores=0.40397686,
final_local_scores=0.3965795, pem=-0.9568352,
similarity=21.127855]
11: 2485308 Gary_Teichmann
    [pem_pos=0, final_scores=0.75404364,
final_local_scores=0.8386288, pem=0.0, similarity=29.11607]
12: 1709031 Andre_Markgraaff
    [pem_pos=0, final_scores=0.4372146,
final_local_scores=0.5082017, pem=0.0, similarity=19.255466]
13: 21833605 Japie_Mulder
    [pem_pos=0, final_scores=1.2129493,
final_local_scores=1.3003383, pem=0.0, similarity=42.894398]
14: 8904211 Northerns_cricket_team
    [pem_pos=2, final_scores=0.29714382,
final_local_scores=0.33681747, pem=-0.52053446,
similarity=16.971766]
15: 8357907 Andre_Snyman
    [pem_pos=0, final_scores=0.6431239,
final_local_scores=0.71000415, pem=0.0, similarity=25.277653]
16: 8357907 Andre_Snyman
```

```
[pem_pos=0, final_scores=0.5923453,
final_local_scores=0.67064375, pem=0.0, similarity=24.10306]
17: 1709031 Andre_Markgraaff
    [pem_pos=0, final_scores=0.4674127,
final_local_scores=0.54118854, pem=0.0, similarity=20.239859]

Sample: b'143_BADMINTON'
BADMINTON - [1 1 MALAYSIAN OPEN 1] BADMINTON RESULTS . [2 2 K
UALA LUMPUR 2] 1996-08-23 Results in the [3 3 Malaysian Open 3
] badminton tournament on Friday ( prefix numbers denote
seedings ) : Men 's singles , quarter-finals 2 - [4 4 Ong Ewe
Hock 4] ( [5 5 Malaysia 5] ) beat 5/8 - Hu Zhilan ( [6 6 China
6] ) 15-2 15-10 9/16 - Luo Yigang ( [7 7 China 7] ) beat
Jason Wong ( [8 8 Malaysia 8] ) 15-5 15-6 Ijaya Indra ( [9 9 I
ndonesia 9] ) beat P. Kantharoopan ( [10 10 Malaysia 10] )
15-6 5-4 9/16 - Chen Gang ( [11 11 China 11] ) beat 9/16 - [12
12 Hermawan Susanto 12] ( [13 13 Indonesia 13] ) 15-9 15-7
1: 10145751 Malaysia_Open_(badminton)
2: 16854 Kuala_Lumpur
3: 10145751 Malaysia_Open_(badminton)
4: 8278861 Ong_Ewe_Hock
5: 3607937 Malaysia
6: 5405 China
7: 5405 China
8: 3607937 Malaysia
9: 14579 Indonesia
10: 3607937 Malaysia
11: 5405 China
12: 9521761 Hermawan_Susanto
13: 14579 Indonesia
1: 10145751 Malaysia_Open_(badminton)
    [pem_pos=2, final_scores=0.42795175,
final_local_scores=0.4902135, pem=-1.3940765,
similarity=26.299866]
2: 16854 Kuala_Lumpur
    [pem_pos=0, final_scores=0.45268613,
final_local_scores=0.43175703, pem=0.0, similarity=16.974203]
3: 10145751 Malaysia_Open_(badminton)
    [pem_pos=2, final_scores=0.38230938,
final_local_scores=0.4403562, pem=-1.3940765,
similarity=24.812025]
4: 8278861 Ong_Ewe_Hock
    [pem_pos=0, final_scores=0.4235559,
final_local_scores=0.5911428, pem=0.0, similarity=21.730593]
5: 3607937 Malaysia
    [pem_pos=0, final_scores=0.64651227,
final_local_scores=0.60540354, pem=-0.20223404,
similarity=23.255941]
6: 5405 China
    [pem_pos=0, final_scores=0.48072308,
final_local_scores=0.46027398, pem=-0.15669827,
similarity=18.677357]
```

```
7: 5405 China
    [pem_pos=0, final_scores=0.48852462,
final_local_scores=0.468796, pem=-0.15669827,
similarity=18.931671]
8: 3607937 Malaysia
    [pem_pos=0, final_scores=0.6677914,
final_local_scores=0.6286478, pem=-0.20223404,
similarity=23.949596]
9: 14579 Indonesia
    [pem_pos=0, final_scores=0.6507516,
final_local_scores=0.62584656, pem=-0.23151328,
similarity=24.025229]
10: 3607937 Malaysia
    [pem_pos=0, final_scores=0.6969277,
final_local_scores=0.6604748, pem=-0.20223404,
similarity=24.899378]
11: 5405 China
    [pem_pos=0, final_scores=0.5021325,
final_local_scores=0.48366052, pem=-0.15669827,
similarity=19.375257]
12: 9521761 Hermawan_Susanto
    [pem_pos=0, final_scores=0.742688,
final_local_scores=0.8605604, pem=0.0, similarity=29.770552]
13: 14579 Indonesia
    [pem_pos=0, final_scores=0.65690917,
final_local_scores=0.6325728, pem=-0.23151328,
similarity=24.225954]
```

Sample: b'144_TENNIS'

TENNIS - INJURED [1 1 CHANDA RUBIN 1] OUT OF [2 2 U.S. OPEN 2] . [3 3 NEW YORK 3] 1996-08-23 Promising 10th-ranked [4 4 American 4] [5 5 Chanda Rubin 5] has pulled out of the [6 6 U.S. Open Tennis Championships 6] with a wrist injury , tournament officials announced . The 20-year-old [7 7 Rubin 7] , who was to be seeded 11th , is still suffering from tendinitis of the right wrist that has kept her sidelined in recent months . [8 8 Rubin 8] 's misfortune turned into a very lucky break for eighth-seeded [9 9 Olympic 9] champion [10 10 Lindsay Davenport 10] . [11 11 Davenport 11] had drawn one of the toughest first-round assignments of any of the seeded players in 17th-ranked [12 12 Karina Habsudova 12] of [13 13 Slovakia 13] . But as the highest-ranked non-seeded player in the tournament , Habsudova will be moved into [14 14 Rubin 14] 's slot in the draw , while [15 15 Davenport 15] will now get a qualifier in the first round , according to [16 16 U.S. Tennis Association 16] officials . [17 17 Rubin 17] is the third notable withdrawal from the women 's competition after 12th-ranked former [18 18 Australian Open 18] champion [19 19 Mary Pierce 19] and 20th-ranked [20 20 Wimbledon 20] semifinalist [21 21 Meredith McGrath 21] pulled out earlier this week with injuries . Men 's [22 22 Australian Open 22] champion [23 23 Boris Becker 23] will also miss the year 's final [24 24 Grand

Slam 24] with a wrist injury .

1: 1496412 Chanda_Rubin
2: 82729 US_Open_(tennis)
3: 645042 New_York_City
4: 3434750 United_States
5: 1496412 Chanda_Rubin
6: 82729 US_Open_(tennis)
7: 1496412 Chanda_Rubin
8: 1496412 Chanda_Rubin
9: 27566 Summer_Olympic_Games
10: 52828 Lindsay_Davenport
11: 52828 Lindsay_Davenport
12: 3062115 Karina_Habšudová
13: 26830 Slovakia
14: 1496412 Chanda_Rubin
15: 52828 Lindsay_Davenport
16: 1178132 United_States_Tennis_Association
17: 1496412 Chanda_Rubin
18: 97851 Australian_Open
19: 754866 Mary_Pierce
20: 91195 The_Championships,_Wimbledon
21: 14979526 Meredith_McGrath
22: 97851 Australian_Open
23: 201049 Boris_Becker
24: 197638 Grand_Slam_(tennis)

1: 1496412 Chanda_Rubin
[pem_pos=0, final_scores=0.48012918,
final_local_scores=0.43776917, pem=0.0, similarity=17.153618]
2: 82729 US_Open_(tennis)
[pem_pos=1, final_scores=0.5456963,
final_local_scores=0.53127664, pem=-1.2370842,
similarity=26.671524]
3: 645042 New_York_City
[pem_pos=1, final_scores=0.336909,
final_local_scores=0.45973235, pem=-1.7906687,
similarity=27.546978]
4: 3434750 United_States
[pem_pos=0, final_scores=0.54285383,
final_local_scores=0.6943073, pem=-0.12994486,
similarity=25.51589]
5: 1496412 Chanda_Rubin
[pem_pos=0, final_scores=0.7839852,
final_local_scores=0.7696857, pem=0.0, similarity=27.058668]
6: 82729 US_Open_(tennis)
[pem_pos=0, final_scores=0.5680863,
final_local_scores=0.5557343, pem=0.0, similarity=20.673933]
7: 1496412 Chanda_Rubin
[pem_pos=0, final_scores=1.0197108,
final_local_scores=1.0306768, pem=0.0, similarity=34.847168]
8: 1496412 Chanda_Rubin
[pem_pos=0, final_scores=1.2219402,
final_local_scores=1.2515817, pem=0.0, similarity=41.439407]

```
9: 27566 Summer_Olympic_Games
    [pem_pos=2, final_scores=0.33958703,
final_local_scores=0.40546507, pem=-3.2532303,
similarity=33.881172]
10: 52828 Lindsay_Davenport
    [pem_pos=0, final_scores=0.9444416,
final_local_scores=0.93837947, pem=0.0, similarity=32.092827]
11: 52828 Lindsay_Davenport
    [pem_pos=0, final_scores=1.2303576,
final_local_scores=1.2536709, pem=0.0, similarity=41.501755]
12: 3062115 Karina_Habšudová
    [pem_pos=0, final_scores=0.76655334,
final_local_scores=0.8590098, pem=0.0, similarity=29.72428]
13: 26830 Slovakia
    [pem_pos=0, final_scores=0.5430895,
final_local_scores=0.6736313, pem=-0.36602598,
similarity=26.182722]
14: 1496412 Chanda_Rubin
    [pem_pos=0, final_scores=1.2113221,
final_local_scores=1.2399831, pem=0.0, similarity=41.09328]
15: 52828 Lindsay_Davenport
    [pem_pos=0, final_scores=1.1871772,
final_local_scores=1.2065029, pem=0.0, similarity=40.09417]
16: 1178132 United_States_Tennis_Association
    [pem_pos=0, final_scores=0.31453532,
final_local_scores=0.33596355, pem=0.0, similarity=14.115536]
17: 1496412 Chanda_Rubin
    [pem_pos=0, final_scores=1.1901062,
final_local_scores=1.2168078, pem=0.0, similarity=40.401688]
18: 97851 Australian_Open
    [pem_pos=0, final_scores=0.4354725,
final_local_scores=0.41124475, pem=-0.3261775,
similarity=18.13588]
19: 754866 Mary_Pierce
    [pem_pos=0, final_scores=0.88831407,
final_local_scores=0.8936706, pem=0.0, similarity=30.758627]
20: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.52520615,
final_local_scores=0.5122602, pem=-0.8505555,
similarity=24.002031]
21: 14979526 Meredith_McGrath
    [pem_pos=0, final_scores=0.7442753,
final_local_scores=0.78837436, pem=0.0, similarity=27.616377]
22: 97851 Australian_Open
    [pem_pos=0, final_scores=0.43678588,
final_local_scores=0.41267943, pem=-0.3261775,
similarity=18.178694]
23: 201049 Boris_Becker
    [pem_pos=0, final_scores=0.74262124,
final_local_scores=0.7492735, pem=0.0, similarity=26.449528]
24: 197638 Grand_Slam_(tennis)
    [pem_pos=0, final_scores=0.47935814,
```

```
final_local_scores=0.45995736, pem=-0.37449613,  
similarity=19.852327]
```

Sample: b'145_BASEBALL'

BASEBALL - [1 1 MAJOR LEAGUE 1] STANDINGS AFTER THURSDAY 'S GAMES . [2 2 NEW YORK 2] 1996-08-23 [3 3 Major League Baseball 3] standings after games played on Thursday (tabulate under won , lost , winning percentage and games behind) : [4 4 AMERICAN LEAGUE EASTERN DIVISION 4] W L PCT GB [5 5 NEW YORK 5] 72 54 .571 - [6 6 BALTIMORE 6] 67 59 .532 5 [7 7 BOSTON 7] 64 64 .500 9 [8 8 TORONTO 8] 59 69 .461 14 [9 9 DETROIT 9] 45 82 .354 27 1/2 [10 10 CENTRAL DIVISION 10] [11 11 CLEVELAND 11] 76 51 .598 - 1 CHICAGO 69 60 .535 8 [12 12 MINNESOTA 12] 63 64 .496 13 [13 13 MILWAUKEE 13] 60 68 .469 16 1/2 [14 14 KANSAS CITY 14] 58 71 .450 19 [15 15 WESTERN DIVISION 15] [16 16 TEXAS 16] 74 54 .578 - [17 17 SEATTLE 17] 65 61 .516 8 [18 18 OAKLAND 18] 62 68 .477 13 2 CALIFORNIA 59 68 .465 14 1/2 FRIDAY , AUGUST 23 SCHEDULE [19 19 SEATTLE 19] AT [20 20 BOSTON 20] [21 21 MILWAUKEE 21] AT [22 22 CLEVELAND 22] 3 CALIFORNIA AT [23 23 BALTIMORE 23] [24 24 OAKLAND 24] AT [25 25 NEW YORK 25] [26 26 TORONTO 26] AT 4 CHICAGO [27 27 DETROIT 27] AT [28 28 KANSAS CITY 28] [29 29 TEXAS 29] AT [30 30 MINNESOTA 30] [31 31 NATIONAL LEAGUE EASTERN DIVISION 31] W L PCT GB [32 32 ATLANTA 32] 79 47 .627 - [33 33 MONTREAL 33] 68 58 .540 11 [34 34 NEW YORK 34] 59 69 .461 21 5 FLORIDA 58 69 .457 21 1/2 [35 35 PHILADELPHIA 35] 52 76 .406 28 [36 36 CENTRAL DIVISION 36] [37 37 HOUSTON 37] 68 60 .531 - ST LOUIS 67 60 .528 1/2 6 CHICAGO 63 62 .504 3 1/2 [38 38 CINCINNATI 38] 63 62 .504 3 1/2 [39 39 PITTSBURGH 39] 54 73 .425 13 1/2 [40 40 WESTERN DIVISION 40] [41 41 SAN DIEGO 41] 70 59 .543 - 7 LOS ANGELES 67 60 .528 2 [42 42 COLORADO 42] 66 62 .516 3 1/2 [43 43 SAN FRANCISCO 43] 54 71 .432 14 FRIDAY , AUGUST 23 SCHEDULE [44 44 CINCINNATI 44] AT 8 FLORIDA (doubleheader) 9 CHICAGO AT [45 45 ATLANTA 45] ST LOUIS AT [46 46 HOUSTON 46] [47 47 PITTSBURGH 47] AT [48 48 COLORADO 48] [49 49 NEW YORK 49] AT 10 LOS ANGELS [50 50 PHILADELPHIA 50] AT [51 51 SAN DIEGO 51] [52 52 MONTREAL 52] AT [53 53 SAN FRANCISCO 53]

1: 38776 Major_League_Baseball
2: 645042 New_York_City
3: 38776 Major_League_Baseball
4: 2004582 American_League_East
5: 4848143 New_York_Yankees
6: 4340 Baltimore_Orioles
7: 4339 Boston_Red_Sox
8: 30847 Toronto_Blue_Jays
9: 8579 Detroit_Tigers
10: 2456377 National_League_Central
11: 6652 Cleveland_Indians
12: 20050 Minnesota_Twins
13: 20152 Milwaukee_Brewers
14: 16883 Kansas_City_Royals
15: 2456474 National_League_West

```
16: 30857 Texas_Rangers_(baseball)
17: 27660 Seattle_Mariners
18: 22523 Oakland_Athletics
19: 27660 Seattle_Mariners
20: 4339 Boston_Red_Sox
21: 20152 Milwaukee_Brewers
22: 6652 Cleveland_Indians
23: 4340 Baltimore_Orioles
24: 22523 Oakland_Athletics
25: 4848143 New_York_Yankees
26: 30847 Toronto_Blue_Jays
27: 8579 Detroit_Tigers
28: 16883 Kansas_City_Royals
29: 30857 Texas_Rangers_(baseball)
30: 20050 Minnesota_Twins
31: 2456294 National_League_East
32: 2140 Atlanta_Braves
33: 20153 Montreal_Expos
34: 4848143 New_York_Yankees
35: 23741 Philadelphia_Phillies
36: 2456377 National_League_Central
37: 13894 Houston_Astros
38: 6671 Cincinnati_Reds
39: 24270 Pittsburgh_Pirates
40: 2456474 National_League_West
41: 28417 San_Diego_Padres
42: 6669 Colorado_Rockies
43: 28416 San_Francisco_Giants
44: 6671 Cincinnati_Reds
45: 2140 Atlanta_Braves
46: 13894 Houston_Astros
47: 24270 Pittsburgh_Pirates
48: 6669 Colorado_Rockies
49: 4848143 New_York_Yankees
50: 23741 Philadelphia_Phillies
51: 28417 San_Diego_Padres
52: 20153 Montreal_Expos
53: 28416 San_Francisco_Giants
1: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.7304415, final_local_scores=-0.81980926,
pem=-4.591283, similarity=4.5930777|
     RECALL MISS: 5945 Chicago_White_Sox]
2: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 7590031 California_Golden_Bears_football,
pem_pos=1, final_scores=-0.8813591,
final_local_scores=-0.96712834, pem=-4.4771376,
similarity=-0.42395434|
     RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
3: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 7590031 California_Golden_Bears_football,
pem_pos=1, final_scores=-0.8377622,
```

```
final_local_scores=-0.9195053, pem=-4.4771376,
similarity=0.9972116|
    RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
4: 5945 Chicago_White_Sox
    [winner: 6886 Chicago, pem_pos=0,
final_scores=-0.6352232, final_local_scores=-0.6878827,
pem=-0.20658566, similarity=-15.314633|
    RECALL MISS: 5945 Chicago_White_Sox]
5: 11278 Miami_Marlins
    [winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.7480026, final_local_scores=-0.81859803,
pem=-4.620281, similarity=4.786922|
    RECALL MISS: 11278 Miami_Marlins]
6: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.75196946, final_local_scores=-0.8433253,
pem=-4.591283, similarity=3.8913116|
    RECALL MISS: 5945 Chicago_White_Sox]
7: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 18213 Los_Angeles_Dodgers, pem_pos=15,
final_scores=-0.17063423, final_local_scores=-0.30179387,
pem=-4.7524934, similarity=20.928381|
    RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
8: 11278 Miami_Marlins
    [winner: 18933066 Florida, pem_pos=0,
final_scores=-0.5762135, final_local_scores=-0.5831213,
pem=-0.15405534, similarity=-12.474012|
    RECALL MISS: 11278 Miami_Marlins]
9: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.6793133, final_local_scores=-0.76395947,
pem=-4.591283, similarity=6.259747|
    RECALL MISS: 5945 Chicago_White_Sox]
10: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 18110 Los_Angeles, pem_pos=0,
final_scores=0.0075002164, final_local_scores=0.033543557,
pem=-0.100652285, similarity=5.6380825|
    RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
1: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.5895823,
final_local_scores=0.542191, pem=-0.47792417,
similarity=22.8688]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.3573897,
final_local_scores=0.42484927, pem=-1.7906687,
similarity=26.505997]
3: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.54041386,
final_local_scores=0.48848194, pem=0.0, similarity=18.666988]
4: 2004582 American_League_East
    [pem_pos=0, final_scores=0.4779967,
final_local_scores=0.47634494, pem=0.0, similarity=18.304796]
```

```
5: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.5012916,
final_local_scores=0.4341529, pem=-5.0692368,
similarity=44.613]
6: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.2818346,
final_local_scores=1.2842095, pem=-4.046561,
similarity=64.41891]
7: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.1238258,
final_local_scores=1.1155939, pem=-4.418931,
similarity=61.412094]
8: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.8543617,
final_local_scores=0.82630885, pem=-4.422637,
similarity=52.799404]
9: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.1994218,
final_local_scores=1.1944375, pem=-4.1401463,
similarity=62.24887]
10: 2456377 National_League_Central
    [pem_pos=0, final_scores=0.57930595,
final_local_scores=0.59783727, pem=-0.98951703,
similarity=27.311516]
11: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.2800096,
final_local_scores=1.2815137, pem=-4.191953,
similarity=65.129135]
12: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=1.2307588,
final_local_scores=1.2310421, pem=-4.2906566,
similarity=64.15972]
13: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.5147395,
final_local_scores=0.45029622, pem=-3.40205,
similarity=36.028328]
14: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.8415565,
final_local_scores=0.80814016, pem=-3.5915844,
similarity=47.737823]
15: 2456474 National_League_West
    [pem_pos=1, final_scores=0.554848,
final_local_scores=0.5575396, pem=-1.2952092,
similarity=27.771355]
16: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.2161846,
final_local_scores=1.2162044, pem=-4.928425,
similarity=67.18522]
17: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.7634912,
final_local_scores=0.7280796, pem=-3.8567436,
similarity=46.79063]
```

```
18: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.2715847,
final_local_scores=1.2736282, pem=-3.9188845,
similarity=63.408825]
19: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.70899004,
final_local_scores=0.6685453, pem=-3.8567436,
similarity=45.014008]
20: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.0223229,
final_local_scores=1.0047174, pem=-4.418931,
similarity=58.103317]
21: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.5158366,
final_local_scores=0.45149463, pem=-3.40205,
similarity=36.06409]
22: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.2170748,
final_local_scores=1.2127669, pem=-4.191953,
similarity=63.077587]
23: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.0432436,
final_local_scores=1.0235851, pem=-4.046561,
similarity=56.641357]
24: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.2214398,
final_local_scores=1.2188525, pem=-3.9188845,
similarity=61.774208]
25: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.27285397,
final_local_scores=0.18486184, pem=-5.0692368,
similarity=37.173656]
26: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.65061146,
final_local_scores=0.6037426, pem=-4.422637,
similarity=46.157585]
27: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.1851959,
final_local_scores=1.1788979, pem=-4.1401463,
similarity=61.785137]
28: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.8012338,
final_local_scores=0.7640939, pem=-3.5915844,
similarity=46.423393]
29: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.2010053,
final_local_scores=1.1996233, pem=-4.928425,
similarity=66.69041]
30: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=0.9362963,
final_local_scores=0.9093866, pem=-4.2906566,
similarity=54.56088]
```

```
31: 2456294 National_League_East
    [pem_pos=0, final_scores=0.43350202,
final_local_scores=0.42720503, pem=0.0, similarity=16.838364]
32: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=1.1332811,
final_local_scores=1.1218104, pem=-4.136474,
similarity=60.061565]
33: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.8104256,
final_local_scores=0.7755929, pem=-4.5098605,
similarity=51.760277]
34: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.56457597,
final_local_scores=0.5032814, pem=-5.0692368,
similarity=46.675934]
35: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.2383709,
final_local_scores=1.237957, pem=-4.5189333,
similarity=65.60748]
36: 2456377 National_League_Central
    [pem_pos=0, final_scores=0.57767195,
final_local_scores=0.59605235, pem=-0.98951703,
similarity=27.258253]
37: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.3424255,
final_local_scores=1.3518095, pem=-4.650927,
similarity=69.72287]
38: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.1425002,
final_local_scores=1.1330559, pem=-3.771247,
similarity=58.410988]
39: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.2876495,
final_local_scores=1.2899661, pem=-4.03702,
similarity=64.53882]
40: 2456474 National_League_West
    [pem_pos=1, final_scores=0.5775723,
final_local_scores=0.5823624, pem=-1.2952092,
similarity=28.512117]
41: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.9228304,
final_local_scores=0.897451, pem=-3.9307637,
similarity=52.247543]
42: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.1195042,
final_local_scores=1.1153772, pem=-4.482383,
similarity=61.750683]
43: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.96649426,
final_local_scores=0.940034, pem=-4.409061,
similarity=56.11936]
44: 6671 Cincinnati_Reds
```

```
[pem_pos=4, final_scores=0.9929324,
final_local_scores=0.96967584, pem=-3.771247,
similarity=53.5354]
45: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=1.0709844,
final_local_scores=1.0537608, pem=-4.136474,
similarity=58.030827]
46: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.1359998,
final_local_scores=1.1263208, pem=-4.650927,
similarity=62.993843]
47: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.3040801,
final_local_scores=1.307914, pem=-4.03702,
similarity=65.07442]
48: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.2771724,
final_local_scores=1.2876058, pem=-4.482383,
similarity=66.89033]
49: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.598042,
final_local_scores=0.539838, pem=-5.0692368,
similarity=47.766853]
50: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.1207823,
final_local_scores=1.1095092, pem=-4.5189333,
similarity=61.77434]
51: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.92150027,
final_local_scores=0.89599806, pem=-3.9307637,
similarity=52.204185]
52: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.6639731,
final_local_scores=0.6156157, pem=-4.5098605,
similarity=46.986237]
53: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.83198315,
final_local_scores=0.793101, pem=-4.409061,
similarity=51.73459]
```

Sample: b'146_BASEBALL'

```
BASEBALL - [1 1 MAJOR LEAGUE 1] RESULTS THURSDAY . [2 2 NEW YORK 2] 1996-08-23 Results of [3 3 Major League 3] Baseball games played on Thursday ( home team in CAPS ) : [4 4 American League 4] [5 5 BOSTON 5] 2 [6 6 Oakland 6] 1 [7 7 Seattle 7] 10 [8 8 BALTIMORE 8] 3 1 California 12 [9 9 NEW YORK 9] 3 [10 10 Toronto 10] 1 2 CHICAGO 0 ( in 6 1/2 ) [11 11 Detroit 11] 10 [12 12 KANSAS CITY 12] 3 [13 13 Texas 13] 11 [14 14 MINNESOTA 14] 2 [15 15 National League 15] [16 16 COLORADO 16] 10 St Louis 5 [17 17 Cincinnati 17] 3 [18 18 ATLANTA 18] 2 ( in 13 ) [19 19 Pittsburgh 19] 8 [20 20 HOUSTON 20] 6 3 LOS ANGELS 8 [21 21 Philadelphia 21] 5 [22 22 Montreal 22] 5 [23 23 SAN
```

```
N FRANCISCO 23] 4
1: 38776 Major_League_Baseball
2: 645042 New_York_City
3: 38776 Major_League_Baseball
4: 64777 American_League
5: 4339 Boston_Red_Sox
6: 22523 Oakland_Athletics
7: 27660 Seattle_Mariners
8: 4340 Baltimore_Orioles
9: 4848143 New_York_Yankees
10: 30847 Toronto_Blue_Jays
11: 8579 Detroit_Tigers
12: 16883 Kansas_City_Royals
13: 30857 Texas_Rangers_(baseball)
14: 20050 Minnesota_Twins
15: 64290 National_League
16: 6669 Colorado_Rockies
17: 6671 Cincinnati_Reds
18: 2140 Atlanta_Braves
19: 24270 Pittsburgh_Pirates
20: 13894 Houston_Astros
21: 23741 Philadelphia_Phillies
22: 20153 Montreal_Expos
23: 28416 San_Francisco_Giants
1: 1360083 Los_Angeles_Angels_of_Anheim
    [winner: 7590031 California_Golden_Bears_football,
pem_pos=1, final_scores=-0.69496745,
final_local_scores=-0.7606303, pem=-4.4771376,
similarity=5.7383575|
     RECALL MISS: 1360083 Los_Angeles_Angels_of_Anheim]
2: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.73315936, final_local_scores=-0.821427,
pem=-4.591283, similarity=4.544802|
     RECALL MISS: 5945 Chicago_White_Sox]
3: 1360083 Los_Angeles_Angels_of_Anheim
    [winner: 18213 Los_Angeles_Dodgers, pem_pos=15,
final_scores=-0.17533617, final_local_scores=-0.3061821,
pem=-4.7524934, similarity=20.797426|
     RECALL MISS: 1360083 Los_Angeles_Angels_of_Anheim]
1: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.55000955,
final_local_scores=0.49550837, pem=-0.47792417,
similarity=21.475697]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.29609203,
final_local_scores=0.36426288, pem=-1.7906687,
similarity=24.697979]
3: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.5603239,
final_local_scores=0.50677526, pem=-0.47792417,
similarity=21.811922]
```

```
4: 64777 American_League
    [pem_pos=0, final_scores=0.53576976,
final_local_scores=0.50251526, pem=0.0, similarity=19.08577]
5: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=0.94328636,
final_local_scores=0.91890794, pem=-4.418931,
similarity=55.54259]
6: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=0.9481428,
final_local_scores=0.9202905, pem=-3.9188845,
similarity=52.86452]
7: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.46784288,
final_local_scores=0.4070186, pem=-3.8567436,
similarity=37.20953]
8: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.0759366,
final_local_scores=1.0585384, pem=-4.046561,
similarity=57.684437]
9: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.48949426,
final_local_scores=0.42285174, pem=-5.0692368,
similarity=44.27575]
10: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.73463935,
final_local_scores=0.697729, pem=-4.422637,
similarity=48.962326]
11: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.0491557,
final_local_scores=1.0295095, pem=-4.1401463,
similarity=57.327095]
12: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.71723425,
final_local_scores=0.67280614, pem=-3.5915844,
similarity=43.699184]
13: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=0.94858164,
final_local_scores=0.92511743, pem=-4.928425,
similarity=58.498604]
14: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=1.0906093,
final_local_scores=1.076689, pem=-4.2906566,
similarity=59.55352]
15: 64290 National_League
    [pem_pos=0, final_scores=0.58781844,
final_local_scores=0.55407697, pem=-0.07431644,
similarity=21.02862]
16: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.0708241,
final_local_scores=1.0652959, pem=-4.482383,
similarity=60.25616]
17: 6671 Cincinnati_Reds
```

```
[pem_pos=4, final_scores=1.0990411,
final_local_scores=1.0851917, pem=-3.771247,
similarity=56.982624]
18: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=0.972019,
final_local_scores=0.9460011, pem=-4.136474,
similarity=54.815063]
19: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.0708907,
final_local_scores=1.0525889, pem=-4.03702,
similarity=57.45501]
20: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.2835265,
final_local_scores=1.2891283, pem=-4.650927,
similarity=67.85234]
21: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.1709465,
final_local_scores=1.163857, pem=-4.5189333,
similarity=63.396187]
22: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.91199535,
final_local_scores=0.8878389, pem=-4.5098605,
similarity=55.10992]
23: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.8563327,
final_local_scores=0.819784, pem=-4.409061,
similarity=52.53086]
```

Sample: b'147_BASEBALL'

BASEBALL - [1 1 SORRENTO 1] HITS SLAM AS [2 2 SEATTLE 2]
ROUTS [3 3 ORIOLES 3] . [4 4 BALTIMORE 4] 1996-08-22 Former [5
5 Oriole 5] [6 6 Jamie Moyer 6] allowed two hits over eight
scoreless innings before tiring in the ninth and [7 7 Paul So
rrento 7] added his third grand slam of the season as the [8 8
Seattle Mariners 8] routed [9 9 Baltimore 9] 10-3 Thursday . [
10 10 Moyer 10] (10-2) , who was tagged for a pair of
homers by [11 11 Mike Devereaux 11] and [12 12 Brady Anderson
12] and three runs in the ninth , walked none and struck out
two . [13 13 Norm Charlton 13] retired the final three
batters to seal the victory . With one out in the fifth [14 14
Ken Griffey Jr 14] and [15 15 Edgar Martinez 15] stroked back-
to-back singles off [16 16 Orioles 16] starter [17 17 Rocky Co
ppinger 17] (7-5) and [18 18 Jay Buhner 18] walked . [19 19
Sorrento 19] followed by hitting a 1-2 pitch just over the
right-field wall for a 7-0 advantage . Right fielder [20 20 Bo
bby Bonilla 20] was after the ball , which was touched by
fans at the top of the scoreboard in right . " Things fell in
for us , " said [21 21 Sorrento 21] , who has six career
grand slams and hit the ninth of the season for the [22 22 Ma
riners 22] . " We have over a month left . We 've got to make
up some ground . " In the [23 23 American League 23] wild-
card race , the [24 24 Mariners 24] are three games behind

the [25 25 White Sox 25] , two behind [26 26 Baltimore 26] and two ahead of the [27 27 Red Sox 27] heading into [28 28 Boston 28] for a weekend series . [29 29 Moyer 29] retired 11 straight batters between the third and seventh innings and threw two or fewer pitches to 11 of the 29 batters he faced . " I made some bad pitches at the end but I 'm not going to dwell on it . We won the game , " said [30 30 Moyer 30] . [31 31 Copper 31] (7-5) was tagged for eight runs and 10 hits in 4 1/3 innings . [32 32 Orioles 32] manager [33 33 Dav ey Johnson 33] missed the game after being admitted to a hospital with an irregular heartbeat . Bench coach [34 34 Andy Etchebarren 34] took his place . In [35 35 Boston 35] , Troy O'Leary homered off the right-field foul pole with one out in the bottom of the ninth and the [36 36 Red Sox 36] climbed to the .500 mark for the first time this season with their fourth straight victory , 2-1 over the [37 37 Oakland Athleti cs 37] . [38 38 Boston 38] has won 15 of its last 19 games . [39 39 Boston 39] 's [40 40 Roger Clemens 40] (7-11) was one out away from his second straight shutout when pinch-hitter [41 41 Matt Stairs 41] tripled over the head of centre fielder [42 42 Lee Tinsley 42] on an 0-2 pitch and pinch-hitter [43 43 Terry Steinbach 43] dunked a broken-bat single into right to lift [44 44 Oakland 44] into a 1-1 tie . The run broke [45 45 Clemens 45] ' 28-inning shutout streak , longest in the majors this season . He pitched his fourth complete game , allowing eight hits with two walks and 11 strikeouts . Reliever [46 46 Mark Acre 46] (0-1) took the loss . In [47 47 New York 47] , [48 48 Garret Anderson 48] and [49 49 Gary Di Sarcina 49] drove in two runs apiece in a five-run first inning and [50 50 Jim Edmonds 50] highlighted a six-run sixth with a bases-loaded double as the [51 51 California Angels 51] coasted to a 12-3 victory over the [52 52 Yankees 52] in the rubber game of their three-game series . The [53 53 Angels 53] battered [54 54 Kenny Rogers 54] (10-7) for five runs in the first . The [55 55 Yankees 55] have allowed at least two runs in the first inning in six straight games , getting outscored 21-1 in the first inning in that span . [56 56 Chuck Finley 56] (12-12) snapped a four-game losing streak . In [57 57 Kansas City 57] , [58 58 Travis Fryman 58] doubled in the go-ahead run in the fifth and [59 59 Melvin Nieves 59] and [60 60 Damion Easley 60] belted two-run homers as the [61 61 Detroit Tigers 61] claimed a 10-3 win over the [62 62 Roya ls 62] , handing them their fifth straight loss . The [63 63 T igers 63] won their third straight and halted a seven-game road losing streak behind [64 64 Justin Thompson 64] (1-2) , who earned his first major-league win . [65 65 Tim B elcher 65] (12-8) was tagged for six runs and nine hits in eight innings . At [66 66 Minnesota 66] , [67 67 Ken Hill 67] allowed two runs en route to his sixth complete game of the season and [68 68 Rusty Greer 68] added three hits , including a homer , and two RBI as the red-hot [69 69 Texas Ra ngers 69] routed the [70 70 Twins 70] 11-2 . The [71 71 Range

rs 71] , who won for the 11th time in their last 13 games , have scored 45 runs in their last five contests . [72 72 Hill 72] (14-7) allowed 10 hits . He has yielded just seven runs in his last four starts , covering 33 1/3 innings . In [73 73 Chicago 73] , [74 74 Erik Hanson 74] outdueled [75 75 Alex Fernandez 75] , and [76 76 Jacob Brumfield 76] drove in [77 77 0 tis Nixon 77] with the game 's only run in the sixth inning as the [78 78 Toronto Blue Jays 78] blanked the [79 79 White Sox 79] 1-0 in a game shortened to six innings due to rain . [8 0 80 Toronto 80] won its fifth straight and handed the [81 81 White Sox 81] their seventh loss in nine games . [82 82 Hanson 82] (11-15) allowed three hits , walked three and struck out four to snap a personal three-game losing streak . Fernandez (12-8) scattered six hits .

1: 3328363 Paul_Sorrento
2: 27660 Seattle_Mariners
3: 4340 Baltimore_Orioles
4: 26997138 Baltimore
5: 4340 Baltimore_Orioles
6: 1661437 Jamie_Moyer
7: 3328363 Paul_Sorrento
8: 27660 Seattle_Mariners
9: 4340 Baltimore_Orioles
10: 1661437 Jamie_Moyer
11: 4440986 Mike_Devereaux
12: 390266 Brady_Anderson
13: 2428651 Norm_Charlton
14: 63493 Ken_Griffey,_Jr.
15: 572954 Edgar_Martínez
16: 4340 Baltimore_Orioles
17: 18939619 Rocky_Coppinger
18: 385914 Jay_Buhner
19: 3328363 Paul_Sorrento
20: 905899 Bobby_Bonilla
21: 3328363 Paul_Sorrento
22: 27660 Seattle_Mariners
23: 64777 American_League
24: 27660 Seattle_Mariners
25: 5945 Chicago_White_Sox
26: 4340 Baltimore_Orioles
27: 4339 Boston_Red_Sox
28: 24437894 Boston
29: 1661437 Jamie_Moyer
30: 1661437 Jamie_Moyer
31: 18939619 Rocky_Coppinger
32: 4340 Baltimore_Orioles
33: 652850 Davey_Johnson
34: 1928809 Andy_Etchebarren
35: 24437894 Boston
36: 4339 Boston_Red_Sox
37: 22523 Oakland_Athletics
38: 4339 Boston_Red_Sox

```
39: 4339 Boston_Red_Sox
40: 26074 Roger_Clemens
41: 88499 Matt_Stairs
42: 5355977 Lee_Tinsley
43: 2806775 Terry_Steinbach
44: 22523 Oakland_Athletics
45: 26074 Roger_Clemens
46: 18370398 Mark_Acre
47: 645042 New_York_City
48: 1458407 Garret_Anderson
49: 2406277 Gary_DiSarcina
50: 881905 Jim_Edmunds
51: 1360083 Los_Angeles_Angels_ofAnaheim
52: 4848143 New_York_Yankees
53: 1360083 Los_Angeles_Angels_ofAnaheim
54: 599018 Kenny_Rogers_(baseball)
55: 4848143 New_York_Yankees
56: 869667 Chuck_Finley
57: 17454 Kansas_City,_Missouri
58: 2406202 Travis_Fryman
59: 9717489 Melvin_Nieves
60: 2406257 Damion_Easley
61: 8579 Detroit_Tigers
62: 16883 Kansas_City_Royals
63: 8579 Detroit_Tigers
64: 2747849 Justin_Thompson
65: 3651486 Tim_Belcher
66: 19590 Minnesota
67: 5085409 Ken_Hill_(baseball)
68: 2070208 Rusty_Greer
69: 30857 Texas_Rangers_(baseball)
70: 20050 Minnesota_Twins
71: 30857 Texas_Rangers_(baseball)
72: 5085409 Ken_Hill_(baseball)
73: 6886 Chicago
74: 2372135 Erik_Hanson_(baseball)
75: 30871773 Alex_Fernandez_(baseball)
76: 10381057 Jacob_Brumfield
77: 1135016 Otis_Nixon
78: 30847 Toronto_Blue_Jays
79: 5945 Chicago_White_Sox
80: 30847 Toronto_Blue_Jays
81: 5945 Chicago_White_Sox
82: 2372135 Erik_Hanson_(baseball)
1: 3328363 Paul_Sorrento
    [pem_pos=0, final_scores=0.87125975,
final_local_scores=0.9274798, pem=0.0, similarity=31.767561]
2: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.72153693,
final_local_scores=0.6841323, pem=-3.8567436,
similarity=45.479156]
3: 4340 Baltimore_Orioles
```

```
[pem_pos=0, final_scores=0.5546318,
final_local_scores=0.49851644, pem=-0.18171716,
similarity=19.954643]
4: 26997138 Baltimore
    [pem_pos=0, final_scores=0.4797756,
final_local_scores=0.5467306, pem=-0.19949137,
similarity=21.490108]
5: 4340 Baltimore_Orioles
    [pem_pos=7, final_scores=0.42471665,
final_local_scores=0.3566038, pem=-2.2027018,
similarity=26.710115]
6: 1661437 Jamie_Moyer
    [pem_pos=0, final_scores=0.7832672,
final_local_scores=0.7857464, pem=0.0, similarity=27.537952]
7: 3328363 Paul_Sorrento
    [pem_pos=0, final_scores=0.7616449,
final_local_scores=0.8052414, pem=0.0, similarity=28.119722]
8: 27660 Seattle_Mariners
    [pem_pos=0, final_scores=0.4883644,
final_local_scores=0.4294266, pem=-0.008242749,
similarity=16.949486]
9: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=0.97056097,
final_local_scores=0.95285577, pem=-4.046561,
similarity=54.530655]
10: 1661437 Jamie_Moyer
    [pem_pos=0, final_scores=0.8326172,
final_local_scores=0.8411851, pem=0.0, similarity=29.192354]
11: 4440986 Mike_Devereaux
    [pem_pos=0, final_scores=0.5505584,
final_local_scores=0.58206844, pem=0.0, similarity=21.459797]
12: 390266 Brady_Anderson
    [pem_pos=0, final_scores=0.6207492,
final_local_scores=0.64679, pem=0.0, similarity=23.391218]
13: 2428651 Norm_Charlton
    [pem_pos=0, final_scores=0.40982705,
final_local_scores=0.4634617, pem=0.0, similarity=17.920334]
14: 63493 Ken_Griffey,_Jr.
    [pem_pos=0, final_scores=0.56329757,
final_local_scores=0.549962, pem=0.0, similarity=20.501675]
15: 572954 Edgar_Martínez
    [pem_pos=0, final_scores=0.5274392,
final_local_scores=0.5460142, pem=0.0, similarity=20.383865]
16: 4340 Baltimore_Orioles
    [pem_pos=0, final_scores=0.597749,
final_local_scores=0.5456154, pem=-0.18171716,
similarity=21.36017]
17: 18939619 Rocky_Coppinger
    [pem_pos=0, final_scores=0.52094686,
final_local_scores=0.6174058, pem=0.0, similarity=22.514332]
18: 385914 Jay_Buhner
    [pem_pos=0, final_scores=0.6222058,
```

```
final_local_scores=0.6446585, pem=0.0, similarity=23.327608]
19: 3328363 Paul_Sorrento
    [pem_pos=0, final_scores=1.0456583,
final_local_scores=1.1179838, pem=0.0, similarity=37.45258]
20: 905899 Bobby_Bonilla
    [pem_pos=0, final_scores=0.8119622,
final_local_scores=0.858204, pem=0.0, similarity=29.700232]
21: 3328363 Paul_Sorrento
    [pem_pos=0, final_scores=1.1376717,
final_local_scores=1.2184944, pem=0.0, similarity=40.452015]
22: 27660 Seattle_Mariners
    [pem_pos=0, final_scores=0.4740649,
final_local_scores=0.41380656, pem=-0.1261296,
similarity=17.12444]
23: 64777 American_League
    [pem_pos=0, final_scores=0.49358743,
final_local_scores=0.46036446, pem=0.0, similarity=17.827908]
24: 27660 Seattle_Mariners
    [pem_pos=0, final_scores=0.44951415,
final_local_scores=0.38698864, pem=-0.1261296,
similarity=16.324139]
25: 5945 Chicago_White_Sox
    [pem_pos=0, final_scores=0.48161072,
final_local_scores=0.4176004, pem=-0.08007053,
similarity=16.987179]
26: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=0.95690924,
final_local_scores=0.9379434, pem=-4.046561,
similarity=54.08564]
27: 4339 Boston_Red_Sox
    [pem_pos=0, final_scores=0.3521095,
final_local_scores=0.2786764, pem=-0.07716359,
similarity=12.825601]
28: 24437894 Boston
    [pem_pos=0, final_scores=0.3431201,
final_local_scores=0.40931523, pem=-0.23164093,
similarity=17.564196]
29: 1661437 Jamie_Moyer
    [pem_pos=0, final_scores=0.8233152,
final_local_scores=0.8310241, pem=0.0, similarity=28.88913]
30: 1661437 Jamie_Moyer
    [pem_pos=0, final_scores=0.6986466,
final_local_scores=0.69484264, pem=0.0, similarity=24.825203]
31: 18939619 Rocky_Coppinger
    [pem_pos=1, final_scores=0.6961306,
final_local_scores=0.8126286, pem=0.0, similarity=28.340172]
32: 4340 Baltimore_Orioles
    [pem_pos=0, final_scores=0.5024091,
final_local_scores=0.44147104, pem=-0.18171716,
similarity=18.252295]
33: 652850 Davey_Johnson
    [pem_pos=0, final_scores=0.59023744,
```

```
final_local_scores=0.5935325, pem=0.0, similarity=21.801907]
34: 1928809 Andy_Etchebarren
    [pem_pos=0, final_scores=0.5389965,
final_local_scores=0.58241266, pem=0.0, similarity=21.470068]
35: 24437894 Boston
    [pem_pos=0, final_scores=0.47050512,
final_local_scores=0.548464, pem=-0.23164093,
similarity=21.716673]
36: 4339 Boston_Red_Sox
    [pem_pos=0, final_scores=0.432778,
final_local_scores=0.36679453, pem=-0.07716359,
similarity=15.455221]
37: 22523 Oakland_Athletics
    [pem_pos=0, final_scores=0.5819337,
final_local_scores=0.5286054, pem=-0.06842491,
similarity=20.236458]
38: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=0.8467172,
final_local_scores=0.81896025, pem=-4.418931,
similarity=52.55995]
39: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=0.84697205,
final_local_scores=0.8192386, pem=-4.418931,
similarity=52.56826]
40: 26074 Roger_Clemens
    [pem_pos=0, final_scores=0.6660616,
final_local_scores=0.6641654, pem=0.0, similarity=23.909733]
41: 88499 Matt_Stairs
    [pem_pos=0, final_scores=0.83163804,
final_local_scores=0.8468385, pem=0.0, similarity=29.361063]
42: 5355977 Lee_Tinsley
    [pem_pos=0, final_scores=0.48338658,
final_local_scores=0.51465935, pem=0.0, similarity=19.448174]
43: 2806775 Terry_Steinbach
    [pem_pos=0, final_scores=0.67084354,
final_local_scores=0.7124659, pem=0.0, similarity=25.351116]
44: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.0790181,
final_local_scores=1.0715947, pem=-3.9188845,
similarity=57.379738]
45: 26074 Roger_Clemens
    [pem_pos=0, final_scores=0.6760265,
final_local_scores=0.6767358, pem=0.0, similarity=24.284859]
46: 18370398 Mark_Acre
    [pem_pos=0, final_scores=0.7499791,
final_local_scores=0.90343446, pem=0.0, similarity=31.049997]
47: 645042 New_York_City
    [pem_pos=1, final_scores=0.44878322,
final_local_scores=0.5323893, pem=-1.7906687,
similarity=29.715206]
48: 1458407 Garret_Anderson
    [pem_pos=0, final_scores=0.7075893,
```

```
final_local_scores=0.72896475, pem=0.0, similarity=25.843475]
49: 2406277 Gary_DiSarcina
    [pem_pos=0, final_scores=0.6578702,
final_local_scores=0.6892634, pem=0.0, similarity=24.658709]
50: 881905 Jim_Edmunds
    [pem_pos=0, final_scores=0.8664685,
final_local_scores=0.8873728, pem=0.0, similarity=30.570686]
51: 1360083 Los_Angeles_Angels_ofAnaheim
    [pem_pos=0, final_scores=0.55942637,
final_local_scores=0.50393367, pem=-0.016987266,
similarity=19.220478]
52: 4848143 New_York_Yankees
    [pem_pos=0, final_scores=0.4788245,
final_local_scores=0.41768664, pem=-0.4294726,
similarity=18.889853]
53: 1360083 Los_Angeles_Angels_ofAnaheim
    [pem_pos=2, final_scores=0.5224674,
final_local_scores=0.4635616, pem=-1.7538311,
similarity=27.460922]
54: 599018 Kenny_Rogers_(baseball)
    [pem_pos=1, final_scores=0.6026239,
final_local_scores=0.6026772, pem=-0.8663494,
similarity=26.786146]
55: 4848143 New_York_Yankees
    [pem_pos=0, final_scores=0.50667137,
final_local_scores=0.44810516, pem=-0.4294726,
similarity=19.797604]
56: 869667 Chuck_Finley
    [pem_pos=0, final_scores=0.69009596,
final_local_scores=0.7261416, pem=0.0, similarity=25.759226]
57: 17454 Kansas_City,_Missouri
    [pem_pos=0, final_scores=0.3277884,
final_local_scores=0.3594488, pem=-0.8046767,
similarity=19.192339]
58: 2406202 Travis_Fryman
    [pem_pos=0, final_scores=0.6862753,
final_local_scores=0.7273145, pem=0.0, similarity=25.794228]
59: 9717489 Melvin_Nieves
    [pem_pos=0, final_scores=0.5992683,
final_local_scores=0.62663615, pem=0.0, similarity=22.789785]
60: 2406257 Damion_Easley
    [pem_pos=0, final_scores=0.67032653,
final_local_scores=0.7024034, pem=0.0, similarity=25.050833]
61: 8579 Detroit_Tigers
    [pem_pos=0, final_scores=0.58298963,
final_local_scores=0.53156906, pem=-0.08904677,
similarity=20.437044]
62: 16883 Kansas_City_Royals
    [pem_pos=0, final_scores=0.57913274,
final_local_scores=0.5264684, pem=-0.31042016,
similarity=21.488693]
63: 8579 Detroit_Tigers
```

```
[pem_pos=0, final_scores=0.61824125,
final_local_scores=0.5700761, pem=-1.1469651,
similarity=27.339296]
64: 2747849 Justin_Thompson
    [pem_pos=0, final_scores=0.6935035,
final_local_scores=0.76131904, pem=0.0, similarity=26.808992]
65: 3651486 Tim_Belcher
    [pem_pos=0, final_scores=0.69627476,
final_local_scores=0.7182076, pem=0.0, similarity=25.52246]
66: 19590 Minnesota
    [pem_pos=0, final_scores=0.44790107,
final_local_scores=0.5424259, pem=-0.21388707,
similarity=21.439934]
67: 5085409 Ken_Hill_(baseball)
    [pem_pos=0, final_scores=0.67384034,
final_local_scores=0.7245125, pem=-0.32208347,
similarity=27.46215]
68: 2070208 Rusty_Greer
    [pem_pos=0, final_scores=0.6895873,
final_local_scores=0.75859654, pem=0.0, similarity=26.727747]
69: 30857 Texas_Rangers_(baseball)
    [pem_pos=0, final_scores=0.571685,
final_local_scores=0.5197028, pem=-0.15069729,
similarity=20.418196]
70: 20050 Minnesota_Twins
    [pem_pos=1, final_scores=0.44898778,
final_local_scores=0.38726288, pem=-1.3860784,
similarity=23.184118]
71: 30857 Texas_Rangers_(baseball)
    [pem_pos=1, final_scores=0.7262283,
final_local_scores=0.68851787, pem=-2.6930556,
similarity=39.281715]
72: 5085409 Ken_Hill_(baseball)
    [pem_pos=0, final_scores=0.7858432,
final_local_scores=0.84685856, pem=-0.32208347,
similarity=31.1132]
73: 6886 Chicago
    [pem_pos=0, final_scores=0.6411871,
final_local_scores=0.71825475, pem=-0.20658566,
similarity=26.647312]
74: 2372135 Erik_Hanson_(baseball)
    [pem_pos=0, final_scores=0.60211354,
final_local_scores=0.6686948, pem=0.0, similarity=24.044899]
75: 30871773 Alex_Fernandez_(baseball)
    [pem_pos=0, final_scores=0.43862647,
final_local_scores=0.46561527, pem=-0.09630868,
similarity=18.508343]
76: 10381057 Jacob_Brumfield
    [pem_pos=0, final_scores=0.39747858,
final_local_scores=0.5109895, pem=0.0, similarity=19.33866]
77: 1135016 Otis_Nixon
    [pem_pos=0, final_scores=0.63038075,
```

```
final_local_scores=0.67581886, pem=0.0, similarity=24.257496]
78: 30847 Toronto_Blue_Jays
    [pem_pos=0, final_scores=0.4249466,
final_local_scores=0.3633697, pem=-0.018220792,
similarity=15.032477]
79: 5945 Chicago_White_Sox
    [pem_pos=0, final_scores=0.47638232,
final_local_scores=0.4118892, pem=-0.08007053,
similarity=16.816744]
80: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.73687786,
final_local_scores=0.70410717, pem=-4.422637,
similarity=49.152664]
81: 5945 Chicago_White_Sox
    [pem_pos=0, final_scores=0.51028824,
final_local_scores=0.44892627, pem=-0.08007053,
similarity=17.922005]
82: 2372135 Erik_Hanson_(baseball)
    [pem_pos=0, final_scores=0.6902164,
final_local_scores=0.76493376, pem=0.0, similarity=26.916862]
```

Sample: b'148_SOCCER'

SOCCE - [1 1 SPORTING 1] START NEW SEASON WITH A WIN . [2 2 LISBON 2] 1996-08-23 [3 3 Sporting 3] 's Luis Miguel Predrosa scored the first goal of the new league season as the Lisbon side cruised to a 3-1 away win over [4 4 SC Espinho 4] on Friday . Predrosa drilled a right-foot shot into the back of the net after 24 minutes to set [5 5 Sporting 5] on the way to victory . Although [6 6 Espinho 6] 's [7 7 Nail Besirovic 7] put the home side back on terms in the 35th minute , [8 8 Sporting 8] quickly restored their lead . [9 9 Jose Luis Vidigal 9] scored in the 38th minute and [10 10 Mustapha Hadji 10] added the third in the 57th . The game was brought forward from Sunday when reigning champions [11 11 Porto 11] and Lisbon rivals [12 12 Benfica 12] play their first games of the season .

```
1: 769618 Sporting_Clube_de_Portugal
2: 18091 Lisbon
3: 769618 Sporting_Clube_de_Portugal
4: 7308787 S.C._Espinho
5: 769618 Sporting_Clube_de_Portugal
6: 7308787 S.C._Espinho
7: 24204394 Nail_Beširović
8: 769618 Sporting_Clube_de_Portugal
9: 5947406 José_Luís_Vidigal
10: 1152430 Mustapha_Hadji
11: 172589 F.C._Porto
12: 574076 S.L._Benfica
1: 769618 Sporting_Clube_de_Portugal
    [pem_pos=0, final_scores=0.5474109,
final_local_scores=0.48507583, pem=-0.12110038,
similarity=19.223906]
```

```
2: 18091 Lisbon
    [pem_pos=0, final_scores=0.54806936,
final_local_scores=0.5998687, pem=-0.09749677,
similarity=22.521194]
3: 769618 Sporting_Clube_de_Portugal
    [pem_pos=0, final_scores=0.40147537,
final_local_scores=0.32566348, pem=-0.12110038,
similarity=14.466723]
4: 7308787 S.C._Espinho
    [pem_pos=1, final_scores=0.3975218,
final_local_scores=0.37905538, pem=-0.07145894,
similarity=15.790087]
5: 769618 Sporting_Clube_de_Portugal
    [pem_pos=0, final_scores=0.41112548,
final_local_scores=0.33620474, pem=-0.12110038,
similarity=14.781296]
6: 7308787 S.C._Espinho
    [pem_pos=3, final_scores=0.42242318,
final_local_scores=0.40625638, pem=-1.0954815,
similarity=22.170612]
7: 24204394 Nail_Beširović
    [pem_pos=0, final_scores=0.4708343,
final_local_scores=0.574918, pem=0.0, similarity=21.246412]
8: 769618 Sporting_Clube_de_Portugal
    [pem_pos=0, final_scores=0.3780386,
final_local_scores=0.30006227, pem=-0.12110038,
similarity=13.702732]
9: 5947406 José_Luís_Vidigal
    [pem_pos=0, final_scores=0.52464825,
final_local_scores=0.56127346, pem=0.0, similarity=20.839231]
10: 1152430 Mustapha_Hadji
    [pem_pos=0, final_scores=0.70366,
final_local_scores=0.7879017, pem=0.0, similarity=27.602272]
11: 172589 F.C._Porto
    [pem_pos=1, final_scores=0.66395575,
final_local_scores=0.628136, pem=-1.1013846,
similarity=28.824045]
12: 574076 S.L._Benfica
    [pem_pos=0, final_scores=0.46747965,
final_local_scores=0.40601408, pem=0.0, similarity=16.205982]
```

Sample: b'149_SOCCER'

```
SOCCER - [1 1 PORTUGUESE 1] FIRST DIVISION RESULT . [1 1 LIS
BON 1] 1996-08-23 Result of a [2 2 Portuguese 2] first
division soccer match on Friday : [3 3 Espinho 3] 1 [4 4 Spor
ting 4] 3
1: 18091 Lisbon
2: 23033 Portugal
3: 7308787 S.C._Espinho
4: 769618 Sporting_Clube_de_Portugal
1: 909545 Primeira_Liga
    [winner: 23033 Portugal, pem_pos=1,
```

```

final_scores=0.28056067, final_local_scores=0.22827947,
pem=-1.127839, similarity=17.035393]
    RECALL MISS: 909545 Primeira_Liga]
1: 18091 Lisbon
    [pem_pos=0, final_scores=0.64754087,
final_local_scores=0.6446043, pem=-0.09749677,
similarity=23.856195]
2: 23033 Portugal
    [pem_pos=1, final_scores=0.5988285,
final_local_scores=0.57593876, pem=-1.127839,
similarity=27.410238]
3: 7308787 S.C._Espinho
    [pem_pos=3, final_scores=0.3755198,
final_local_scores=0.4156518, pem=-1.0954815,
similarity=22.45099]
4: 769618 Sporting_Clube_de_Portugal
    [pem_pos=0, final_scores=0.26141727,
final_local_scores=0.25086468, pem=-0.12110038,
similarity=12.234578]
1: 23033 Portugal
    [pem_pos=1, final_scores=0.28056067,
final_local_scores=0.22827947, pem=-1.127839,
similarity=17.035393]

Sample: b'150_SOCCER'
SOCCER - [1 1 ST PAULI 1] TAKE POINT WITH LATE FIGHTBACK . [2
2 BONN 2] 1996-08-23 [3 3 Hamburg 3] side [4 4 St Pauli 4] ,
tipped as prime candidates for relegation , produced a
stunning second-half fightback to draw 4-4 in their [5 5 Bund
esliga 5] clash with [6 6 Schalke 6] on Friday . [7 7 Schalke
7] , who finished third last season , raced to a 3-1 lead at
halftime . [8 8 St Pauli 8] pulled a goal back through [9 9 An
dre Trulsen 9] but [10 10 Schalke 10] striker [11 11 Martin Ma
x 11] restored his team 's two-goal cushion shortly
afterwards . [12 12 Christian Springer 12] put [13 13 St Pauli
13] back in touch in the 64th minute and three minutes later
they were level , thanks to a penalty from Thomas Sabotzik .
In the night 's only other match , [14 14 Hamburg 14] beat [15
15 Hansa Rostock 15] 1-0 , [16 16 Karsten Baeron 16] scoring
the winner after some dazzling build-up from in-form
midfielder [17 17 Harald Spoerl 17] . The win put [18 18 Hamb
urg 18] in second place in the [19 19 German 19] first
division after three games , though that may change after the
other sides play on Saturday .
1: 881244 FC_St._Pauli
2: 3295 Bonn
3: 13467 Hamburg
4: 881244 FC_St._Pauli
5: 686142 Bundesliga
6: 410348 FC_Schalke_04
7: 410348 FC_Schalke_04
8: 881244 FC_St._Pauli

```

```
9: 20978193 André_Trulsen
10: 410348 FC_Schalke_04
11: 5208200 Martin_Max
12: 21093872 Christian_Springer
13: 881244 FC_St._Pauli
14: 535879 Hamburger_SV
15: 599978 F.C._Hansa_Rostock
16: 20996083 Karsten_Bäron
17: 20948640 Harald_Spörl
18: 535879 Hamburger_SV
19: 11867 Germany
1: 881244 FC_St._Pauli
    [pem_pos=1, final_scores=0.52703494,
final_local_scores=0.48233962, pem=-0.11686432,
similarity=19.119215]
2: 3295 Bonn
    [pem_pos=0, final_scores=0.38062096,
final_local_scores=0.43124628, pem=0.0, similarity=16.958961]
3: 13467 Hamburg
    [pem_pos=0, final_scores=0.44567585,
final_local_scores=0.48235363, pem=-0.27202022,
similarity=19.963396]
4: 881244 FC_St._Pauli
    [pem_pos=1, final_scores=0.55348617,
final_local_scores=0.5112336, pem=-0.11686432,
similarity=19.98147]
5: 686142 Bundesliga
    [pem_pos=0, final_scores=0.41076964,
final_local_scores=0.3580111, pem=-0.08715299,
similarity=15.24743]
6: 410348 FC_Schalke_04
    [pem_pos=0, final_scores=0.60980755,
final_local_scores=0.5718607, pem=0.0, similarity=21.155176]
7: 410348 FC_Schalke_04
    [pem_pos=0, final_scores=0.4415723,
final_local_scores=0.38808924, pem=0.0, similarity=15.67107]
8: 881244 FC_St._Pauli
    [pem_pos=1, final_scores=0.60387784,
final_local_scores=0.5662789, pem=-0.11686432,
similarity=21.62413]
9: 20978193 André_Trulsen
    [pem_pos=0, final_scores=0.54833645,
final_local_scores=0.65014285, pem=0.0, similarity=23.491272]
10: 410348 FC_Schalke_04
    [pem_pos=0, final_scores=0.50031126,
final_local_scores=0.45225263, pem=0.0, similarity=17.585833]
11: 5208200 Martin_Max
    [pem_pos=0, final_scores=0.59761655,
final_local_scores=0.6675304, pem=0.0, similarity=24.010153]
12: 21093872 Christian_Springer
    [pem_pos=0, final_scores=0.57821465,
final_local_scores=0.69519544, pem=0.0, similarity=24.835732]
```

```
13: 881244 FC_St._Pauli
    [pem_pos=1, final_scores=0.57924646,
final_local_scores=0.53937286, pem=-0.11686432,
similarity=20.821201]
14: 535879 Hamburger_SV
    [pem_pos=1, final_scores=0.7258348,
final_local_scores=0.7034904, pem=-2.820485,
similarity=40.421505]
15: 599978 F.C._Hansa_Rostock
    [pem_pos=0, final_scores=0.5546008,
final_local_scores=0.52606237, pem=0.0, similarity=19.788464]
16: 20996083 Karsten_Bäron
    [pem_pos=0, final_scores=0.6826203,
final_local_scores=0.79633975, pem=0.0, similarity=27.85408]
17: 20948640 Harald_Spörl
    [pem_pos=0, final_scores=0.56110394,
final_local_scores=0.66958904, pem=0.0, similarity=24.071585]
18: 535879 Hamburger_SV
    [pem_pos=1, final_scores=0.7452273,
final_local_scores=0.72467375, pem=-2.820485,
similarity=41.05366]
19: 11867 Germany
    [pem_pos=1, final_scores=0.77593666,
final_local_scores=0.83948153, pem=-1.0465823,
similarity=34.832996]
```

Sample: b'151_ATHLETICS'

ATHLETICS - [1 1 SALAH HISSOU 1] BREAKS 10,000 METRES WORLD RECORD . [2 2 BRUSSELS 2] 1996-08-23 [3 3 Morocco 3] 's [4 4 Salah Hissou 4] broke the men 's 10,000 metres world record on Friday when he clocked 26 minutes 38.08 seconds at the [5 5 Brussels 5] grand prix on Friday . The previous mark of 26:43.53 was set by [6 6 Ethiopia 6] 's [7 7 Haile Gebreselassie 7] in the [8 8 Dutch 8] town of [9 9 Hengelo 9] in June last year .

```
1: 3580285 Salah_Hissou
2: 3708 Brussels
3: 19291 Morocco
4: 3580285 Salah_Hissou
5: 3708 Brussels
6: 187749 Ethiopia
7: 616517 Haile_Gebrselassie
8: 21148 Netherlands
9: 118459 Hengelo
1: 3580285 Salah_Hissou
    [pem_pos=0, final_scores=0.46234322,
final_local_scores=0.5525281, pem=0.0, similarity=20.578253]
2: 3708 Brussels
    [pem_pos=0, final_scores=0.54387623,
final_local_scores=0.547745, pem=-0.025440974,
similarity=20.573868]
3: 19291 Morocco
```

```

        [pem_pos=0, final_scores=0.4815908,
final_local_scores=0.49942362, pem=-0.29598162,
similarity=20.603104]
4: 3580285 Salah_Hissou
        [pem_pos=0, final_scores=0.5812356,
final_local_scores=0.6823999, pem=0.0, similarity=24.453888]
5: 3708 Brussels
        [pem_pos=0, final_scores=0.59420794,
final_local_scores=0.6027248, pem=-0.025440974,
similarity=22.214575]
6: 187749 Ethiopia
        [pem_pos=0, final_scores=0.59657645,
final_local_scores=0.65552586, pem=-0.23532115,
similarity=24.931625]
7: 616517 Haile_Gebrselassie
        [pem_pos=0, final_scores=0.8714431,
final_local_scores=0.97581583, pem=0.0, similarity=33.210003]
8: 21148 Netherlands
        [pem_pos=0, final_scores=0.5630186,
final_local_scores=0.57431895, pem=-0.7943452,
similarity=25.54831]
9: 118459 Hengelo
        [pem_pos=0, final_scores=0.51353765,
final_local_scores=0.5647654, pem=0.0, similarity=20.94344]

```

Sample: b'152_SOCCER'

SOCCER - [1 1 GERMAN 1] FIRST DIVISION SUMMARIES . [2 2 BONN
2] 1996-08-23 Summaries of [3 3 Bundesliga 3] matches on
Friday : [4 4 Hansa Rostock 4] 0 [5 5 Hamburg 5] 1 (Baeron
64th min) . Halftime 0-0 . Attendance 18,500 . [6 6 St Pauli
6] 4 ([7 7 Driller 7] 15th , Trulsen 54th , 1 Springer
64th , [8 8 Sobotzik 8] 67th penalty) [9 9 Schalke 9] 4 (2 M
ax 11th , [10 10 Thon 10] 34th , [11 11 Wilmots 11] 38th ,
Springer 64th) . 1-3 . 19,775 .

1: 11867 Germany
2: 3295 Bonn
3: 686142 Bundesliga
4: 599978 F.C._Hansa_Rostock
5: 535879 Hamburger_SV
6: 881244 FC_St._Pauli
7: 20928096 Martin_Driller
8: 17177695 Thomas_Sobotzik
9: 410348 FC_Schalke_04
10: 4909208 Olaf_Thon
11: 1386277 Marc_Wilmots
1: 21093872 Christian_Springer
[winner: 4866959 Springer_(Transformers), pem_pos=0,
final_scores=-0.4377172, final_local_scores=-0.27549365,
pem=-0.8289672, similarity=0.37648532]
RECALL MISS: 21093872 Christian_Springer]

2: 5208200 Martin_Max
[winner: 1674202 Max_Warburg, pem_pos=11,

```

final_scores=-0.625184, final_local_scores=-0.58366364,
pem=-3.8646593, similarity=7.6886444|
    RECALL MISS: 5208200 Martin_Max]
1: 11867 Germany
    [pem_pos=1, final_scores=0.78763646,
final_local_scores=0.8613185, pem=-1.0465823,
similarity=35.484653]
2: 3295 Bonn
    [pem_pos=0, final_scores=0.40345,
final_local_scores=0.46074402, pem=0.0, similarity=17.839233]
3: 686142 Bundesliga
    [pem_pos=0, final_scores=0.33683485,
final_local_scores=0.28004596, pem=-0.08715299,
similarity=12.9207945]
4: 599978 F.C._Hansa_Rostock
    [pem_pos=0, final_scores=0.54500633,
final_local_scores=0.52185696, pem=0.0, similarity=19.662966]
5: 535879 Hamburger_SV
    [pem_pos=1, final_scores=0.79211694,
final_local_scores=0.77926874, pem=-2.820485,
similarity=42.682884]
6: 881244 FC_St._Pauli
    [pem_pos=1, final_scores=0.5695992,
final_local_scores=0.54870754, pem=-0.11686432,
similarity=21.099768]
7: 20928096 Martin_Driller
    [pem_pos=1, final_scores=0.44211978,
final_local_scores=0.5704178, pem=-0.6931472,
similarity=24.88156]
8: 17177695 Thomas_Sobotzik
    [pem_pos=0, final_scores=0.4673273,
final_local_scores=0.5379713, pem=0.0, similarity=20.14385]
9: 410348 FC_Schalke_04
    [pem_pos=0, final_scores=0.5107607,
final_local_scores=0.47198486, pem=0.0, similarity=18.174683]
10: 4909208 Olaf_Thon
    [pem_pos=0, final_scores=0.6660407,
final_local_scores=0.73727566, pem=-0.3051171,
similarity=27.750765]
11: 1386277 Marc_Wilmots
    [pem_pos=0, final_scores=0.4527045,
final_local_scores=0.5156067, pem=0.0, similarity=19.476446]

```

Sample: b'153_SOCCER'

SOCCER - [1 1 FRENCH 1] LEAGUE SUMMARY . [2 2 PARIS 2]

1996-08-23 Summary of a [3 3 French 3] first division match
on Friday . [4 4 Nancy 4] 0 [5 5 Paris St Germain 5] 0 .
Attendance : 15,000 .

1: 5843419 France
2: 22989 Paris
3: 5843419 France
4: 1640855 AS_Nancy

```
5: 357488 Paris_Saint-Germain_F.C.
1: 5843419 France
    [pem_pos=0, final_scores=0.73010534,
final_local_scores=0.7239527, pem=-0.80223274,
similarity=30.056574]
2: 22989 Paris
    [pem_pos=0, final_scores=0.57657844,
final_local_scores=0.57661664, pem=-0.115301244,
similarity=21.92413]
3: 5843419 France
    [pem_pos=0, final_scores=0.815193,
final_local_scores=0.81689817, pem=-0.80223274,
similarity=32.83025]
4: 1640855 AS_Nancy
    [pem_pos=0, final_scores=0.68003803,
final_local_scores=0.7337206, pem=-1.6882259,
similarity=35.166233]
5: 357488 Paris_Saint-Germain_F.C.
    [pem_pos=0, final_scores=0.46262902,
final_local_scores=0.483234, pem=0.0, similarity=18.51038]

Sample: b'154_SOCCER'
SOCCER - [1 1 FRENCH 1] LEAGUE RESULT . [2 2 PARIS 2]
1996-08-23 Result of a [3 3 French 3] first division match on
Friday . [4 4 Nancy 4] 0 [5 5 Paris St Germain 5] 0
1: 5843419 France
2: 22989 Paris
3: 5843419 France
4: 1640855 AS_Nancy
5: 357488 Paris_Saint-Germain_F.C.
1: 5843419 France
    [pem_pos=0, final_scores=0.74283284,
final_local_scores=0.7378556, pem=-0.80223274,
similarity=30.471464]
2: 22989 Paris
    [pem_pos=0, final_scores=0.5684463,
final_local_scores=0.56773347, pem=-0.115301244,
similarity=21.659039]
3: 5843419 France
    [pem_pos=0, final_scores=0.79413134,
final_local_scores=0.7938915, pem=-0.80223274,
similarity=32.143684]
4: 1640855 AS_Nancy
    [pem_pos=0, final_scores=0.62837666,
final_local_scores=0.67728835, pem=-1.6882259,
similarity=33.482185]
5: 357488 Paris_Saint-Germain_F.C.
    [pem_pos=0, final_scores=0.44719213,
final_local_scores=0.4663716, pem=0.0, similarity=18.007172]

Sample: b'155_SOCCER'
SOCCER - [1 1 GERMAN 1] FIRST DIVISION RESULTS . [2 2 BONN 2]
```

```
1996-08-23 Results of [3 3 German 3] first division soccer
matches on Friday : [4 4 St Pauli 4] 4 [5 5 Schalke 5] 4 [6 6
Hansa Rostock 6] 0 [7 7 Hamburg 7] 1
1: 11867 Germany
2: 3295 Bonn
3: 11867 Germany
4: 881244 FC_St._Pauli
5: 410348 FC_Schalke_04
6: 599978 F.C._Hansa_Rostock
7: 535879 Hamburger_SV
1: 11867 Germany
    [pem_pos=1, final_scores=0.84594053,
final_local_scores=0.8737393, pem=-1.0465823,
similarity=35.855316]
2: 3295 Bonn
    [pem_pos=0, final_scores=0.44533747,
final_local_scores=0.46725094, pem=0.0, similarity=18.033413]
3: 11867 Germany
    [pem_pos=1, final_scores=0.790965,
final_local_scores=0.8136869, pem=-1.0465823,
similarity=34.063232]
4: 881244 FC_St._Pauli
    [pem_pos=1, final_scores=0.53475726,
final_local_scores=0.53037703, pem=-0.11686432,
similarity=20.552748]
5: 410348 FC_Schalke_04
    [pem_pos=0, final_scores=0.51223314,
final_local_scores=0.4957419, pem=0.0, similarity=18.88364]
6: 599978 F.C._Hansa_Rostock
    [pem_pos=0, final_scores=0.5357279,
final_local_scores=0.53081053, pem=0.0, similarity=19.930159]
7: 535879 Hamburger_SV
    [pem_pos=1, final_scores=0.8042249,
final_local_scores=0.8162205, pem=-2.820485,
similarity=43.7856]
```

Sample: b'156 ATHLETICS'

ATHLETICS - [1 1 Masterkova 1] BREAKS SECOND WORLD RECORD IN
10 DAYS . Adrian Warner [2 2 BRUSSELS 2] 1996-08-23 [3 3 Russia 3] 's double [4 4 Olympic 4] champion [5 5 Svetlana Masterkova 5] smashed her second world record in just 10 days on Friday when she bettered the mark for the women 's 1,000 metres . After breaking the world record for the women 's mile in [6 6 Zurich 6] last Wednesday , the [7 7 Olympic 7] 800 and 1,500 metres champion clocked two minutes 28.98 seconds over 1,000 at the [8 8 Brussels 8] grand prix meeting . The Russian ate up the ground in a swift last lap to shave 0.36 seconds off the previous best of 2:29.34 set by [9 9 Mozambique 9] 's [10 10 Maria Mutola 10] in the same stadium in August last year . Former world 800 champion [11 11 Mutola 11] pushed [12 12 Masterkova 12] all the way , finishing second in 2:29.66 . But it was the Russian who

picked up the bonus of \$ 25,000 for the historic run in front of a capacity 40,000 crowd . [13 13 Masterkova 13] dominated the middle-distance races at the recent [14 14 Atlanta Games 1 4] following her return to competition this season after a three-year maternity break . In her first mile race at the richest meeting in [15 15 Zurich 15] last Wednesday , she slashed 3.05 seconds off the previous record . The record of four minutes , 12.56 seconds in [16 16 Zurich 16] earned [17 1 7 Masterkova 17] a bonus of \$ 50,000 plus one kilo of gold . After Friday 's performance the Russian will have earned well over \$ 100,000 in less than a fortnight , taking her appearance money into account . [18 18 Brussels 18] organisers had laid a new track for the meeting comparable to the surface at the [19 19 Atlanta Games 19] but put down on a softer surface . [20 20 Masterkova 20] clearly enjoyed it . [2 1 21 Mutola 21] looked threatening in the final 200 metres but the Russian found an extra gear to power home several strides ahead , pointing at the time on the clock with delight as she crossed the line .

```
1: 613736 Svetlana_Masterkova
2: 3708 Brussels
3: 25391 Russia
4: 27566 Summer_Olympic_Games
5: 613736 Svetlana_Masterkova
6: 40334603 Zürich
7: 27566 Summer_Olympic_Games
8: 3708 Brussels
9: 19301 Mozambique
10: 66748 Maria_Mutola
11: 66748 Maria_Mutola
12: 613736 Svetlana_Masterkova
13: 613736 Svetlana_Masterkova
14: 72311 1996_Summer_Olympics
15: 40334603 Zürich
16: 40334603 Zürich
17: 613736 Svetlana_Masterkova
18: 3708 Brussels
19: 72311 1996_Summer_Olympics
20: 613736 Svetlana_Masterkova
21: 66748 Maria_Mutola
1: 613736 Svetlana_Masterkova
    [pem_pos=0, final_scores=0.7996401,
final_local_scores=0.80870605, pem=0.0, similarity=28.223114]
2: 3708 Brussels
    [pem_pos=0, final_scores=0.4859681,
final_local_scores=0.5494176, pem=-0.025440974,
similarity=20.623783]
3: 25391 Russia
    [pem_pos=0, final_scores=0.68411785,
final_local_scores=0.76587397, pem=-0.20299491,
similarity=28.048838]
4: 27566 Summer_Olympic_Games
```

```
[pem_pos=2, final_scores=0.4420522,
final_local_scores=0.4431061, pem=-3.2532303,
similarity=35.004456]
5: 613736 Svetlana_Masterkova
    [pem_pos=0, final_scores=0.9056172,
final_local_scores=0.918593, pem=0.0, similarity=31.50236]
6: 40334603 Zürich
    [pem_pos=0, final_scores=0.57295305,
final_local_scores=0.6408151, pem=-0.08839822,
similarity=23.693636]
7: 27566 Summer_Olympic_Games
    [pem_pos=3, final_scores=0.17883559,
final_local_scores=0.15558201, pem=-3.2532303,
similarity=26.424164]
8: 3708 Brussels
    [pem_pos=0, final_scores=0.48444617,
final_local_scores=0.54775506, pem=-0.025440974,
similarity=20.57417]
9: 19301 Mozambique
    [pem_pos=0, final_scores=0.504973,
final_local_scores=0.62608576, pem=-0.36112452,
similarity=24.737211]
10: 66748 Maria_Mutola
    [pem_pos=0, final_scores=0.69479686,
final_local_scores=0.73094666, pem=0.0, similarity=25.902618]
11: 66748 Maria_Mutola
    [pem_pos=0, final_scores=0.7223206,
final_local_scores=0.7610122, pem=0.0, similarity=26.799835]
12: 613736 Svetlana_Masterkova
    [pem_pos=0, final_scores=0.9970915,
final_local_scores=1.0243917, pem=0.0, similarity=34.659603]
13: 613736 Svetlana_Masterkova
    [pem_pos=0, final_scores=1.0479133,
final_local_scores=1.079907, pem=0.0, similarity=36.31629]
14: 72311 1996_Summer_Olympics
    [pem_pos=0, final_scores=0.3470289,
final_local_scores=0.33794555, pem=0.0, similarity=14.174683]
15: 40334603 Zürich
    [pem_pos=0, final_scores=0.5613099,
final_local_scores=0.6280967, pem=-0.08839822,
similarity=23.314095]
16: 40334603 Zürich
    [pem_pos=0, final_scores=0.57452565,
final_local_scores=0.64253294, pem=-0.08839822,
similarity=23.7449]
17: 613736 Svetlana_Masterkova
    [pem_pos=0, final_scores=0.8903906,
final_local_scores=0.9078372, pem=0.0, similarity=31.181385]
18: 3708 Brussels
    [pem_pos=0, final_scores=0.48611534,
final_local_scores=0.5495784, pem=-0.025440974,
similarity=20.628582]
```

```
19: 72311 1996_Summer_Olympics
    [pem_pos=0, final_scores=0.33365768,
final_local_scores=0.3233395, pem=0.0, similarity=13.738809]
20: 613736 Svetlana_Masterkova
    [pem_pos=0, final_scores=1.034557,
final_local_scores=1.0653172, pem=0.0, similarity=35.880905]
21: 66748 Maria_Mutola
    [pem_pos=0, final_scores=0.6730618,
final_local_scores=0.70720434, pem=0.0, similarity=25.194101]
```

Sample: b'157 ATHLETICS'

ATHLETICS - WOMEN 'S 1,000 METRES WORLD RECORD EVOLUTION . [1
1 BRUSSELS 1] 1996-08-23 Evolution of the women 's 1,000
metres world record (tabulated under time , name /
nationality , venue , date) : 2:30.67 [2 2 Christine Wachtel
2] ([3 3 Germany 3]) [4 4 Berlin 4] 17.8.90 2:29.34 [5 5 Ma
ria Mutola 5] ([6 6 Mozambique 6]) [7 7 Brussels 7] 25.8.95
2:28.98 [8 8 Svetlana Masterkova 8] ([9 9 Russia 9]) [10 10
Brussels 10] 23.8.96

1: 3708 Brussels
2: 4214597 Christine_Wachtel
3: 11867 Germany
4: 3354 Berlin
5: 66748 Maria_Mutola
6: 19301 Mozambique
7: 3708 Brussels
8: 613736 Svetlana_Masterkova
9: 25391 Russia
10: 3708 Brussels
1: 3708 Brussels
 [pem_pos=0, final_scores=0.5814898,
final_local_scores=0.5639947, pem=-0.025440974,
similarity=21.058792]
2: 4214597 Christine_Wachtel
 [pem_pos=0, final_scores=0.47432774,
final_local_scores=0.6251902, pem=0.0, similarity=22.746635]
3: 11867 Germany
 [pem_pos=0, final_scores=0.8513897,
final_local_scores=0.86671823, pem=-0.15529114,
similarity=30.798811]
4: 3354 Berlin
 [pem_pos=0, final_scores=0.6472795,
final_local_scores=0.6329575, pem=-0.15483057,
similarity=23.820421]
5: 66748 Maria_Mutola
 [pem_pos=0, final_scores=0.6689057,
final_local_scores=0.7720744, pem=0.0, similarity=27.129953]
6: 19301 Mozambique
 [pem_pos=0, final_scores=0.5721656,
final_local_scores=0.64781356, pem=-0.36112452,
similarity=25.385612]
7: 3708 Brussels

```
[pem_pos=0, final_scores=0.5872003,
final_local_scores=0.57023257, pem=-0.025440974,
similarity=21.244944]
8: 613736 Svetlana_Masterkova
    [pem_pos=0, final_scores=0.6556373,
final_local_scores=0.806862, pem=0.0, similarity=28.168085]
9: 25391 Russia
    [pem_pos=0, final_scores=0.8626316,
final_local_scores=0.92352456, pem=-0.20299491,
similarity=32.753445]
10: 3708 Brussels
    [pem_pos=0, final_scores=0.5926533,
final_local_scores=0.5761891, pem=-0.025440974,
similarity=21.422697]

Sample: b'158 ATHLETICS'
ATHLETICS - [1 1 MASTERKOVA 1] BREAKS WOMEN 'S WORLD 1,000
RECORD . [2 2 BRUSSELS 2] 1996-08-23 [3 3 Russian 3] [4 4 Sve
tlana Masterkova 4] broke the women 's world 1,000 metres
record on Friday when she clocked an unofficial two minutes
28.99 seconds at the [5 5 Brussels 5] grand prix . The
previous mark of 2:29.34 was set by [6 6 Mozambique 6] 's [7 7
Maria Mutola 7] here on August 25 last year . The time was
officially adjusted to 2:28.98 .
1: 613736 Svetlana_Masterkova
2: 3708 Brussels
3: 25391 Russia
4: 613736 Svetlana_Masterkova
5: 3708 Brussels
6: 19301 Mozambique
7: 66748 Maria_Mutola
1: 613736 Svetlana_Masterkova
    [pem_pos=0, final_scores=0.73100275,
final_local_scores=0.82116526, pem=0.0, similarity=28.594923]
2: 3708 Brussels
    [pem_pos=0, final_scores=0.5384472,
final_local_scores=0.59920746, pem=-0.025440974,
similarity=22.109612]
3: 25391 Russia
    [pem_pos=1, final_scores=0.785407,
final_local_scores=0.86774784, pem=-1.0515022,
similarity=35.703274]
4: 613736 Svetlana_Masterkova
    [pem_pos=0, final_scores=0.7436364,
final_local_scores=0.79133415, pem=0.0, similarity=27.704702]
5: 3708 Brussels
    [pem_pos=0, final_scores=0.53490233,
final_local_scores=0.59533525, pem=-0.025440974,
similarity=21.994057]
6: 19301 Mozambique
    [pem_pos=0, final_scores=0.5640167,
final_local_scores=0.6734654, pem=-0.36112452,
```

```
similarity=26.151114]
7: 66748 Maria_Mutola
    [pem_pos=0, final_scores=0.56324166,
final_local_scores=0.6201953, pem=0.0, similarity=22.597578]
```

Sample: b'159_GOLF'

GOLF - [1 1 GERMAN OPEN 1] SECOND ROUND SCORES . [2 2 STUTTGART 2] , [3 3 Germany 3] 1996-08-23 Leading second round scores in the [4 4 German Open 4] golf championship on Friday ([5 5 Britain 5] unless stated) : 128 [6 6 Ian Woosnam 6] 64 64 129 [7 7 Robert Karlsson 7] ([8 8 Sweden 8]) 67 62 130 Fernando Roca ([9 9 Spain 9]) 66 64 , [10 10 Ian Pyman 10] 66 64 131 [11 11 Carl Suneson 11] 65 66 , Stephen Field 66 65 132 [12 12 Miguel Angel Martin 12] ([13 13 Spain 13]) 66 66 , [14 14 Raymond Russell 14] 63 69 , Thomas Gogele ([15 15 Germany 15]) 67 65 , [16 16 Paul Broadhurst 16] 62 70 , [17 17 Diego Borrego 17] ([18 18 Spain 18]) 69 63 133 Ricky Willison 69 64 , [19 19 Stephen Ames 19] (Trinidad and Tobago) 68 65 , [20 20 Eamonn Darcy 20] ([21 21 Ireland 21]) 65 68 134 [22 22 Robert Coles 22] 68 66 , David Williams 67 67 , [23 23 Thomas Bjorn 23] ([24 24 Denmark 24]) 66 68 , [25 25 Pedro Linhart 25] ([26 26 Spain 26]) 67 67 , [27 27 Michael Jonzon 27] ([28 28 Sweden 28]) 67 67 , [29 29 Roger Chapman 29] 72 62 , [30 30 Jonathan Lomas 30] 67 67 , Francisco Cea ([31 31 Spain 31]) 68 66 135 [32 32 Terry Price 32] ([33 33 Australia 33]) 67 68 , [34 34 Paul Eales 34] 67 68 , [35 35 Wayne Riley 35] ([36 36 Australia 36]) 64 71 , [37 37 Carl Mason 37] 69 66 , [38 38 Barry Lane 38] 68 67 , [39 39 Bernhard Langer 39] ([40 40 Germany 40]) 64 71 , [41 41 Gary Orr 41] 67 68 , [42 42 Mats Lanner 42] ([43 43 Sweden 43]) 64 71 , Jeff Hawksworth 67 68 , [44 44 Des Smyth 44] ([45 45 Ireland 45]) 66 69 , [46 46 David Carter 46] 66 69 , [47 47 Steve Webster 47] 69 66 , [48 48 Jose Maria Canizares 48] ([49 49 Spain 49]) 67 68 , [50 50 Paul Lawrie 50] 66 69

1: 4488686 German_Open_(golf)

2: 28565 Stuttgart

3: 11867 Germany

4: 4488686 German_Open_(golf)

5: 31717 United_Kingdom

6: 1034991 Ian_Woosnam

7: 1988802 Robert_Karlsson

8: 5058739 Sweden

9: 26667 Spain

10: 5088382 Iain_Pyman

11: 11818919 Carl_Suneson

12: 5022198 Miguel_Ángel_Martín_(golfer)

13: 26667 Spain

14: 4375906 Raymond_Russell

15: 11867 Germany

16: 1688190 Paul_Broadhurst

17: 5271765 Diego_Borrego

18: 26667 Spain

```
19: 780871 Stephen_Ames
20: 2770292 Eamonn_Darcy
21: 14560 Republic_of_Ireland
22: 23115502 Robert_Coles_(golfer)
23: 1891067 Thomas_Bjørn
24: 76972 Denmark
25: 5285076 Pedro_Linhart
26: 26667 Spain
27: 5242970 Michael_Jonzon
28: 5058739 Sweden
29: 4380250 Roger_Chapman_(golfer)
30: 5272617 Jonathan_Lomas
31: 26667 Spain
32: 22856557 Terry_Price_(golfer)
33: 4689264 Australia
34: 4435993 Paul_Eales
35: 5243170 Wayne_Riley
36: 4689264 Australia
37: 5222898 Carl_Mason
38: 2167839 Barry_Lane
39: 575815 Bernhard_Langer
40: 11867 Germany
41: 2167716 Gary_Orr
42: 4919505 Mats_Lanner
43: 5058739 Sweden
44: 5193762 Des_Smyth
45: 14560 Republic_of_Ireland
46: 4344819 David_Carter_(golfer)
47: 4375908 Steve_Webster
48: 4455122 José_Maria_Cañizares
49: 26667 Spain
50: 1564436 Paul_Lawrie
1: 4488686 German_Open_(golf)
    [pem_pos=0, final_scores=0.6140041,
final_local_scores=0.6844158, pem=-0.62093085,
similarity=27.890764]
2: 28565 Stuttgart
    [pem_pos=0, final_scores=0.5193418,
final_local_scores=0.58673507, pem=-0.15141529,
similarity=22.422478]
3: 11867 Germany
    [pem_pos=0, final_scores=0.7612966,
final_local_scores=0.8224839, pem=-0.15529114,
similarity=29.478771]
4: 4488686 German_Open_(golf)
    [pem_pos=0, final_scores=0.4480576,
final_local_scores=0.5031445, pem=-0.62093085,
similarity=22.481266]
5: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5653818,
final_local_scores=0.62258416, pem=-0.61114657,
similarity=25.992376]
```

6: 1034991 Ian_Woosnam
[pem_pos=0, final_scores=0.63934046,
final_local_scores=0.674338, pem=0.0, similarity=24.213303]
7: 1988802 Robert_Karlsson
[pem_pos=0, final_scores=0.782101,
final_local_scores=0.86855525, pem=0.0, similarity=30.009132]
8: 5058739 Sweden
[pem_pos=0, final_scores=0.70210654,
final_local_scores=0.7427487, pem=-0.2135508,
similarity=27.416138]
9: 26667 Spain
[pem_pos=0, final_scores=0.6246382,
final_local_scores=0.657622, pem=-0.18857051,
similarity=24.73994]
10: 5088382 Iain_Pyman
[pem_pos=0, final_scores=0.72403866,
final_local_scores=0.8052987, pem=0.0, similarity=28.121431]
11: 11818919 Carl_Suneson
[pem_pos=0, final_scores=0.7660672,
final_local_scores=0.8344622, pem=0.0, similarity=28.99173]
12: 5022198 Miguel_Ángel_Martín_(golfer)
[pem_pos=0, final_scores=0.73201716,
final_local_scores=0.80146724, pem=0.0, similarity=28.007093]
13: 26667 Spain
[pem_pos=0, final_scores=0.5983074,
final_local_scores=0.6288596, pem=-0.18857051,
similarity=23.881613]
14: 4375906 Raymond_Russell
[pem_pos=0, final_scores=0.61337197,
final_local_scores=0.6658728, pem=0.0, similarity=23.960686]
15: 11867 Germany
[pem_pos=0, final_scores=0.84466356,
final_local_scores=0.9135496, pem=-0.15529114,
similarity=32.196354]
16: 1688190 Paul_Broadhurst
[pem_pos=0, final_scores=0.87149745,
final_local_scores=0.9534994, pem=0.0, similarity=32.544037]
17: 5271765 Diego_Borrego
[pem_pos=0, final_scores=0.44314772,
final_local_scores=0.5177185, pem=0.0, similarity=19.539465]
18: 26667 Spain
[pem_pos=0, final_scores=0.59598935,
final_local_scores=0.62632746, pem=-0.18857051,
similarity=23.80605]
19: 780871 Stephen_Ames
[pem_pos=0, final_scores=0.58031344,
final_local_scores=0.63409483, pem=0.0, similarity=23.012367]
20: 2770292 Eamonn_Darcy
[pem_pos=0, final_scores=0.5257644,
final_local_scores=0.58207023, pem=0.0, similarity=21.45985]
21: 14560 Republic_of_Ireland
[pem_pos=1, final_scores=0.5737837,

```
final_local_scores=0.63116854, pem=-1.5040284,
similarity=31.104183]
22: 23115502 Robert_Coles_(golfer)
    [pem_pos=1, final_scores=0.42926234,
final_local_scores=0.53304785, pem=-0.61510444,
similarity=23.341957]
23: 1891067 Thomas_Bjørn
    [pem_pos=0, final_scores=0.7234807,
final_local_scores=0.77739793, pem=0.0, similarity=27.288818]
24: 76972 Denmark
    [pem_pos=0, final_scores=0.58287007,
final_local_scores=0.6224655, pem=-0.26762357,
similarity=24.120703]
25: 5285076 Pedro_Linhart
    [pem_pos=0, final_scores=0.76387817,
final_local_scores=0.84792554, pem=0.0, similarity=29.393503]
26: 26667 Spain
    [pem_pos=0, final_scores=0.62395674,
final_local_scores=0.65687764, pem=-0.18857051,
similarity=24.717728]
27: 5242970 Michael_Jonzon
    [pem_pos=0, final_scores=0.6956188,
final_local_scores=0.77997285, pem=0.0, similarity=27.365658]
28: 5058739 Sweden
    [pem_pos=0, final_scores=0.7177649,
final_local_scores=0.759853, pem=-0.2135508,
similarity=27.926563]
29: 4380250 Roger_Chapman_(golfer)
    [pem_pos=1, final_scores=0.582398,
final_local_scores=0.63092417, pem=-0.5573004,
similarity=25.948435]
30: 5272617 Jonathan_Lomas
    [pem_pos=0, final_scores=0.7714469,
final_local_scores=0.8541974, pem=0.0, similarity=29.580667]
31: 26667 Spain
    [pem_pos=0, final_scores=0.59531397,
final_local_scores=0.6255897, pem=-0.18857051,
similarity=23.784035]
32: 22856557 Terry_Price_(golfer)
    [pem_pos=1, final_scores=0.5718192,
final_local_scores=0.64975995, pem=-0.47743857,
similarity=26.076231]
33: 4689264 Australia
    [pem_pos=0, final_scores=0.6737689,
final_local_scores=0.7424683, pem=-0.17168477,
similarity=27.180098]
34: 4435993 Paul_Eales
    [pem_pos=0, final_scores=0.7236446,
final_local_scores=0.7979111, pem=0.0, similarity=27.900972]
35: 5243170 Wayne_Riley
    [pem_pos=0, final_scores=0.6309537,
final_local_scores=0.68304306, pem=0.0, similarity=24.47308]
```

36: 4689264 Australia
[pem_pos=0, final_scores=0.64125377,
final_local_scores=0.7069505, pem=-0.17168477,
similarity=26.120174]
37: 5222898 Carl_Mason
[pem_pos=0, final_scores=0.76759225,
final_local_scores=0.842174, pem=0.0, similarity=29.221865]
38: 2167839 Barry_Lane
[pem_pos=0, final_scores=0.6405512,
final_local_scores=0.7027933, pem=0.0, similarity=25.062468]
39: 575815 Bernhard_Langer
[pem_pos=0, final_scores=0.9415919,
final_local_scores=1.0057592, pem=0.0, similarity=34.103573]
40: 11867 Germany
[pem_pos=0, final_scores=0.8356194,
final_local_scores=0.90367025, pem=-0.15529114,
similarity=31.901535]
41: 2167716 Gary_Orr
[pem_pos=0, final_scores=0.73213017,
final_local_scores=0.81998694, pem=0.0, similarity=28.55976]
42: 4919505 Mats_Lanner
[pem_pos=0, final_scores=0.8396029,
final_local_scores=0.91051835, pem=0.0, similarity=31.261398]
43: 5058739 Sweden
[pem_pos=0, final_scores=0.7711143,
final_local_scores=0.8181291, pem=-0.2135508,
similarity=29.66564]
44: 5193762 Des_Smyth
[pem_pos=0, final_scores=0.71067685,
final_local_scores=0.77665675, pem=0.0, similarity=27.266699]
45: 14560 Republic_of_Ireland
[pem_pos=1, final_scores=0.5192817,
final_local_scores=0.5716334, pem=-1.5040284,
similarity=29.327536]
46: 4344819 David_Carter_(golfer)
[pem_pos=4, final_scores=0.5434763,
final_local_scores=0.5883822, pem=-1.8641758,
similarity=31.785887]
47: 4375908 Steve_Webster
[pem_pos=0, final_scores=0.5396141,
final_local_scores=0.5934252, pem=0.0, similarity=21.798704]
48: 4455122 José_Maria_Cañizares
[pem_pos=0, final_scores=0.63086,
final_local_scores=0.6904812, pem=0.0, similarity=24.69505]
49: 26667 Spain
[pem_pos=0, final_scores=0.6349846,
final_local_scores=0.6689239, pem=-0.18857051,
similarity=25.077213]
50: 1564436 Paul_Lawrie
[pem_pos=0, final_scores=0.7106462,
final_local_scores=0.7796385, pem=0.0, similarity=27.355682]

Sample: b'160_ATHLETICS'

ATHLETICS - [1 1 MITCHELL 1] UPSTAGES TRIO OF 1 OLYMPIC SPRINT CHAMPIONS . Adrian Warner [2 2 BRUSSELS 2] 1996-08-23 [3 3 American 3] [4 4 Dennis Mitchell 4] upstaged a trio of past and present [5 5 Olympic 5] 100 metres champions on Friday with a storming victory at the [6 6 Brussels 6] grand prix . Sporting his customary bright green outfit , the [7 7 U.S. 7] champion clocked 10.03 seconds despite damp conditions to take the scalp of [8 8 Canada 8] 's reigning [9 9 Olympic 9] champion [10 10 Donovan Bailey 10] , 1992 champion [11 11 Linford Christie 11] of [12 12 Britain 12] and [13 13 American 13] 1984 and 1988 champion [14 14 Carl Lewis 14] . [15 15 Mitchell 15] also beat world and [16 16 Olympic 16] champion [17 17 Bailey 17] at the most lucrative meeting in the sport in [18 18 Zurich 18] last week . The American , who finished fourth at the [19 19 Atlanta Games 19] , was fast out of his blocks and held off [20 20 Bailey 20] 's late burst in the final 20 metres before heading off for a lap of celebration . The Canadian was second in 10.09 with [21 21 Lewis 21] third in 10.10 , ahead of [22 22 Atlanta 22] bronze medallist [23 23 Antonio Boldon 23] who clocked 10.12 in fourth . [24 24 Christie 24] , competing in what is expected to be his last major international meeting , finished fifth in 10.14 . [25 25 Lewis 25] , making a rare appearance in [26 26 Europe 26] in a sprint race , left the track with a slight limp . American Olympic high hurdles champion [27 27 Allen Johnson 27] defied the wet conditions to produce a brilliant 12.92 seconds in the 110 metres race , just 0.01 outside the world record held by [28 28 Britain 28] 's [29 29 Colin Jackson 29] . [30 30 Johnson 30] ran the same time at the [31 31 U.S. 31] [32 32 Olympic 32] trials in [33 33 Atlanta 33] in June to become the second equal fastest hurdler of all time with [34 34 American 34] [35 35 Roger Kingdom 35] . He seemed to relish the new track at the [36 36 Brussels 36] meeting , dominating the race from start to finish with a slight wind at his back . [37 37 Jackson 37] , the only man to have run faster , could not live with his speed , taking second in 13.24 seconds . The rain was pelting down when the women 's high hurdlers stepped up for their 100 metres race . But [38 38 Sweden 38] 's [39 39 Olympic 39] high hurdles champion [40 40 Ludmila Engquist 40] , who crashed out of last week 's meeting in [41 41 Zurich 41] after hitting a hurdle , also kept her footing perfectly to win in a fast 12.60 seconds . [42 42 Olympic 42] silver medallist [43 43 Brigitte Bukovec 43] of [44 44 Slovenia 44] could finish only fifth in the race in 12.95 . [45 45 Jamaican 45] [46 46 Commonwealth 46] champion [47 47 Michelle Freeman 47] took second in 12.77 ahead of [48 48 Cuban 48] [49 49 Aliuskas Lopez 49] . The [50 50 Zurich 50] fall cost [51 51 Engquist 51] a shot at a jackpot of 20 one-kg gold bars which can be won by athletes who clinch their events at all of the Golden Four series in [52 52 Oslo 52] , [53 53 Zurich 53] , [54 54 Brussels 54] and [55 55 Berlin 55] . Seven athletes went

into Friday 's penultimate meeting of the series with a chance of winning the prize and [56 56 American 56] men ' s 400 metres hurdles champion [57 57 Derrick Adkins 57] kept his hopes alive in the competition by winning his event in 47.93 . American Olympic champion [58 58 Gail Devers 58] clocked a swift 10.84 seconds on her way to victory in the women ' s 100 metres , the second fastest time of the season and 0.10 seconds faster than her winning time in [59 59 Atlanta 59] . [60 60 Jamaican 60] veteran [61 61 Merlene Ottey 61] , who beat [62 62 Devers 62] in [63 63 Zurich 63] after just missing out on the gold medal in [64 64 Atlanta 64] after a photo finish , had to settle for third place in 11.04 . [65 65 American 65] world champion [66 66 Gwen Torrence 66] , the bronze medallist in [67 67 Atlanta 67] , was second in 11.00 . It was a costly defeat for [68 68 Ottey 68] since it threw her out of the race for the Golden Four jackpot .

1: 1148663 Dennis_Mitchell
2: 3708 Brussels
3: 3434750 United_States
4: 1148663 Dennis_Mitchell
5: 27566 Summer_Olympic_Games
6: 3708 Brussels
7: 3434750 United_States
8: 5042916 Canada
9: 27566 Summer_Olympic_Games
10: 67116 Donovan_Bailey
11: 67118 Linford_Christie
12: 31717 United_Kingdom
13: 3434750 United_States
14: 67127 Carl_Lewis
15: 1148663 Dennis_Mitchell
16: 27566 Summer_Olympic_Games
17: 67116 Donovan_Bailey
18: 40334603 Zürich
19: 72311 1996_Summer_Olympics
20: 67116 Donovan_Bailey
21: 67127 Carl_Lewis
22: 72311 1996_Summer_Olympics
23: 96607 Ato_Boldon
24: 67118 Linford_Christie
25: 67127 Carl_Lewis
26: 9239 Europe
27: 802066 Allen_Johnson
28: 31717 United_Kingdom
29: 789006 Colin_Jackson
30: 802066 Allen_Johnson
31: 3434750 United_States
32: 27566 Summer_Olympic_Games
33: 3138 Atlanta
34: 3434750 United_States
35: 804243 Roger_Kingdom
36: 3708 Brussels

```
37: 789006 Colin_Jackson
38: 5058739 Sweden
39: 27566 Summer_Olympic_Games
40: 30873921 Ludmila_Engquist
41: 40334603 Zürich
42: 27566 Summer_Olympic_Games
43: 4273494 Brigit_Bukovec
44: 27338 Slovenia
45: 15660 Jamaica
46: 39173 Commonwealth_Games
47: 6829072 Michelle_Freeman
48: 5042481 Cuba
49: 3020050 Aliuska_López
50: 40334603 Zürich
51: 30873921 Ludmila_Engquist
52: 22309 Oslo
53: 40334603 Zürich
54: 3708 Brussels
55: 3354 Berlin
56: 3434750 United_States
57: 993078 Derrick_Adkins
58: 72654 Gail_Devers
59: 72311 1996_Summer_Olympics
60: 15660 Jamaica
61: 558060 Merlene_Otley
62: 72654 Gail_Devers
63: 40334603 Zürich
64: 72311 1996_Summer_Olympics
65: 3434750 United_States
66: 882492 Gwen_Torrence
67: 72311 1996_Summer_Olympics
68: 558060 Merlene_Otley
1: 72311 1996_Summer_Olympics
    [winner: 22576 Olympic_Games, pem_pos=0,
final_scores=-0.04718195, final_local_scores=-0.13837275,
pem=-0.38020664, similarity=2.0280168]
    RECALL MISS: 72311 1996_Summer_Olympics]
1: 1148663 Dennis_Mitchell
    [pem_pos=0, final_scores=0.66752344,
final_local_scores=0.7370641, pem=0.0, similarity=26.085176]
2: 3708 Brussels
    [pem_pos=0, final_scores=0.54484683,
final_local_scores=0.57897633, pem=-0.025440974,
similarity=21.505875]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.5547699,
final_local_scores=0.5650287, pem=-0.12994486,
similarity=21.657959]
4: 1148663 Dennis_Mitchell
    [pem_pos=0, final_scores=0.602651,
final_local_scores=0.6662008, pem=0.0, similarity=23.970474]
5: 27566 Summer_Olympic_Games
```

```
[pem_pos=2, final_scores=0.40208262,
final_local_scores=0.3739676, pem=-3.2532303,
similarity=32.941223]
6: 3708 Brussels
[pem_pos=0, final_scores=0.5847843,
final_local_scores=0.622602, pem=-0.025440974,
similarity=22.80775]
7: 3434750 United_States
[pem_pos=0, final_scores=0.57372874,
final_local_scores=0.58573836, pem=-0.14075686,
similarity=22.334772]
8: 5042916 Canada
[pem_pos=0, final_scores=0.8242306,
final_local_scores=0.8672759, pem=-0.11662793,
similarity=30.605198]
9: 27566 Summer_Olympic_Games
[pem_pos=2, final_scores=0.39257723,
final_local_scores=0.36358434, pem=-3.2532303,
similarity=32.631367]
10: 67116 Donovan_Bailey
[pem_pos=0, final_scores=0.59008896,
final_local_scores=0.64753175, pem=0.0, similarity=23.413351]
11: 67118 Linford_Christie
[pem_pos=0, final_scores=0.960067,
final_local_scores=1.0300348, pem=0.0, similarity=34.828007]
12: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5530362,
final_local_scores=0.5715247, pem=-0.61114657,
similarity=24.46866]
13: 3434750 United_States
[pem_pos=0, final_scores=0.47842306,
final_local_scores=0.4816314, pem=-0.12994486,
similarity=19.169216]
14: 67127 Carl_Lewis
[pem_pos=0, final_scores=0.6188554,
final_local_scores=0.6449707, pem=0.0, similarity=23.336926]
15: 1148663 Dennis_Mitchell
[pem_pos=0, final_scores=0.9082543,
final_local_scores=1.000026, pem=0.0, similarity=33.93248]
16: 27566 Summer_Olympic_Games
[pem_pos=2, final_scores=0.43816864,
final_local_scores=0.41338605, pem=-3.2532303,
similarity=34.11755]
17: 67116 Donovan_Bailey
[pem_pos=0, final_scores=0.7772935,
final_local_scores=0.85202426, pem=0.0, similarity=29.515816]
18: 40334603 Zürich
[pem_pos=0, final_scores=0.5079842,
final_local_scores=0.53256136, pem=-0.08839822,
similarity=20.463129]
19: 72311 1996_Summer_Olympics
[pem_pos=0, final_scores=0.3860131,
```

```
final_local_scores=0.3545281, pem=0.0, similarity=14.6695385]
20: 67116 Donovan_Bailey
    [pem_pos=0, final_scores=0.65692794,
final_local_scores=0.72054315, pem=0.0, similarity=25.592157]
21: 67127 Carl_Lewis
    [pem_pos=0, final_scores=0.7402628,
final_local_scores=0.77758986, pem=0.0, similarity=27.294546]
22: 72311 1996_Summer_Olympics
    [pem_pos=14, final_scores=0.5606637,
final_local_scores=0.54530734, pem=-4.337497,
similarity=43.950756]
23: 96607 Ato_Boldon
    [pem_pos=0, final_scores=0.8560137,
final_local_scores=0.9199658, pem=0.0, similarity=31.54333]
24: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.6926928,
final_local_scores=1.8340781, pem=0.0, similarity=58.822258]
25: 67127 Carl_Lewis
    [pem_pos=0, final_scores=0.7043674,
final_local_scores=0.7383796, pem=0.0, similarity=26.124434]
26: 9239 Europe
    [pem_pos=0, final_scores=0.71829885,
final_local_scores=0.79017496, pem=-0.086338006,
similarity=28.13963]
27: 802066 Allen_Johnson
    [pem_pos=0, final_scores=0.6559249,
final_local_scores=0.7280083, pem=0.0, similarity=25.814932]
28: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5996719,
final_local_scores=0.6224671, pem=-0.61114657,
similarity=25.988884]
29: 789006 Colin_Jackson
    [pem_pos=0, final_scores=0.58122724,
final_local_scores=0.6118367, pem=0.0, similarity=22.348139]
30: 802066 Allen_Johnson
    [pem_pos=0, final_scores=0.63809156,
final_local_scores=0.708528, pem=0.0, similarity=25.2336]
31: 3434750 United_States
    [pem_pos=0, final_scores=0.3926381,
final_local_scores=0.38792437, pem=-0.14075686,
similarity=16.431608]
32: 27566 Summer_Olympic_Games
    [pem_pos=2, final_scores=0.2313088,
final_local_scores=0.18742311, pem=-3.2532303,
similarity=27.374365]
33: 3138 Atlanta
    [pem_pos=0, final_scores=0.5150458,
final_local_scores=0.49969727, pem=-0.23152597,
similarity=20.26075]
34: 3434750 United_States
    [pem_pos=0, final_scores=0.65702975,
final_local_scores=0.67673206, pem=-0.12994486,
```

```
similarity=24.99141]
35: 804243 Roger_Kingdom
    [pem_pos=0, final_scores=0.58331156,
final_local_scores=0.66988134, pem=0.0, similarity=24.080309]
36: 3708 Brussels
    [pem_pos=0, final_scores=0.50915855,
final_local_scores=0.5399924, pem=-0.025440974,
similarity=20.342516]
37: 789006 Colin_Jackson
    [pem_pos=0, final_scores=0.7282911,
final_local_scores=0.77248156, pem=0.0, similarity=27.142103]
38: 5058739 Sweden
    [pem_pos=0, final_scores=0.69208246,
final_local_scores=0.73738766, pem=-0.2135508,
similarity=27.256153]
39: 27566 Summer_Olympic_Games
    [pem_pos=2, final_scores=0.38697225,
final_local_scores=0.35746175, pem=-3.2532303,
similarity=32.448658]
40: 30873921 Ludmila_Engquist
    [pem_pos=0, final_scores=0.6366889,
final_local_scores=0.7518074, pem=0.0, similarity=26.525145]
41: 40334603 Zürich
    [pem_pos=0, final_scores=0.54637307,
final_local_scores=0.5744954, pem=-0.08839822,
similarity=21.714523]
42: 27566 Summer_Olympic_Games
    [pem_pos=2, final_scores=0.26389307,
final_local_scores=0.22301644, pem=-3.2532303,
similarity=28.43654]
43: 4273494 Brigitा_Bukovec
    [pem_pos=0, final_scores=0.7249342,
final_local_scores=0.8808418, pem=0.0, similarity=30.37579]
44: 27338 Slovenia
    [pem_pos=0, final_scores=0.48413694,
final_local_scores=0.5431321, pem=-0.27845925,
similarity=21.812162]
45: 15660 Jamaica
    [pem_pos=0, final_scores=0.43424833,
final_local_scores=0.4408133, pem=-0.14069559,
similarity=18.009584]
46: 39173 Commonwealth_Games
    [pem_pos=10, final_scores=0.26602334,
final_local_scores=0.23549747, pem=-3.517583,
similarity=30.24659]
47: 6829072 Michelle_Freeman
    [pem_pos=0, final_scores=0.6659053,
final_local_scores=0.8030275, pem=0.0, similarity=28.053656]
48: 5042481 Cuba
    [pem_pos=0, final_scores=0.5829597,
final_local_scores=0.61623335, pem=-0.37537703,
similarity=24.520704]
```

49: 3020050 Aliuska_López
[pem_pos=0, final_scores=0.49281603,
final_local_scores=0.64169884, pem=0.0, similarity=23.239286]
50: 40334603 Zürich
[pem_pos=0, final_scores=0.54003966,
final_local_scores=0.56757706, pem=-0.08839822,
similarity=21.508068]
51: 30873921 Ludmila_Engquist
[pem_pos=0, final_scores=0.48502952,
final_local_scores=0.5861426, pem=0.0, similarity=21.581377]
52: 22309 Oslo
[pem_pos=0, final_scores=0.57390237,
final_local_scores=0.61541706, pem=-0.035297178,
similarity=22.646936]
53: 40334603 Zürich
[pem_pos=0, final_scores=0.62267196,
final_local_scores=0.6578403, pem=-0.08839822,
similarity=24.201702]
54: 3708 Brussels
[pem_pos=0, final_scores=0.631517,
final_local_scores=0.6736504, pem=-0.025440974,
similarity=24.331137]
55: 3354 Berlin
[pem_pos=0, final_scores=0.6471579,
final_local_scores=0.6719313, pem=-0.15483057,
similarity=24.983477]
56: 3434750 United_States
[pem_pos=0, final_scores=0.62871253,
final_local_scores=0.6457998, pem=-0.12994486,
similarity=24.068329]
57: 993078 Derrick_Adkins
[pem_pos=0, final_scores=0.61801434,
final_local_scores=0.71407187, pem=0.0, similarity=25.399042]
58: 72654 Gail_Devers
[pem_pos=0, final_scores=0.8191075,
final_local_scores=0.9022375, pem=0.0, similarity=31.014278]
59: 72311 1996_Summer_Olympics
[pem_pos=14, final_scores=0.5246323,
final_local_scores=0.5059485, pem=-4.337497,
similarity=42.77621]
60: 15660 Jamaica
[pem_pos=0, final_scores=0.40412182,
final_local_scores=0.40790468, pem=-0.14069559,
similarity=17.027527]
61: 558060 Merlene_Ottey
[pem_pos=0, final_scores=1.0863129,
final_local_scores=1.1660144, pem=0.0, similarity=38.885906]
62: 72654 Gail_Devers
[pem_pos=0, final_scores=0.773771,
final_local_scores=0.8571843, pem=0.0, similarity=29.669802]
63: 40334603 Zürich
[pem_pos=0, final_scores=0.43861252,

```
final_local_scores=0.4567833, pem=-0.08839822,  
similarity=18.20176]  
64: 72311 1996_Summer_Olympics  
    [pem_pos=14, final_scores=0.47424692,  
final_local_scores=0.45091015, pem=-4.337497,  
similarity=41.133755]  
65: 3434750 United_States  
    [pem_pos=0, final_scores=0.33540732,  
final_local_scores=0.32540843, pem=-0.12994486,  
similarity=14.50721]  
66: 882492 Gwen_Torrence  
    [pem_pos=0, final_scores=0.89377147,  
final_local_scores=1.0095096, pem=0.0, similarity=34.215492]  
67: 72311 1996_Summer_Olympics  
    [pem_pos=14, final_scores=0.51942986,  
final_local_scores=0.50026566, pem=-4.337497,  
similarity=42.606625]  
68: 558060 Merlene_Ottey  
    [pem_pos=0, final_scores=0.66445655,  
final_local_scores=0.70890623, pem=0.0, similarity=25.244888]
```

Sample: b'161 ATHLETICS'
ATHLETICS - BRUSSELS GRAND PRIX RESULTS . [1 1 BRUSSELS 1]
1996-08-23 Leading results in the Brussels Grand Prix
athletics meeting on Friday : Women 's discus 1. [2 2 Ilke Wy
ludda 2] ([3 3 Germany 3]) 66.60 metres 2. [4 4 Ellina Zver
eva 4] ([5 5 Belarus 5]) 65.66 3. [6 6 Franka Dietzsch 6] ([7 7 Germany 7]) 61.74 4. [8 8 Natalya Sadova 8] ([9 9 Russ
ia 9]) 61.64 5. [10 10 Mette Bergmann 10] ([11 11 Norway 11]) 61.44 6. [12 12 Nicoleta Grasu 12] ([13 13 Romania 13]) 61.36 7. [14 14 Olga Chernyavskaya 14] ([15 15 Russia 15]) 60.46 8. [16 16 Irina Yatchenko 16] ([17 17 Belarus 17]) 58.92 Women 's 100 metres hurdles 1. [18 18 Ludmila Engquist 18] ([19 19 Sweden 19]) 12.60 seconds 2. [20 20 Michelle Fre
eman 20] ([21 21 Jamaica 21]) 12.77 3. [22 22 Aliuska Lopez
22] ([23 23 Cuba 23]) 12.85 4. Dionne Rose ([24 24 Jamaica
24]) 12.88 5. [25 25 Brigita Bukovec 25] ([26 26 Slovakia 26]) 12.95 6. Yulia Graudin ([27 27 Russia 27]) 12.96 7. Julie Baumann ([28 28 Switzerland 28]) 13.36 8. [29 29 Patr
icia Girard-Leno 29] ([30 30 France 30]) 13.36 9. Dawn Bowles ([31 31 U.S. 31]) 13.53 Men 's 110 metres hurdles 1. [32 32 Allen Johnson 32] ([33 33 U.S. 33]) 12.92 seconds 2. [34 34 Colin Jackson 34] ([35 35 Britain 35]) 13.24 3. [36 3
6 Emilio Valle 36] ([37 37 Cuba 37]) 13.33 4. [38 38 Sven Pi
eters 38] ([39 39 Belgium 39]) 13.37 5. 1 Steve Brown ([40
40 U.S. 40]) 13.38 6. Frank Asselman ([41 41 Belgium 41]) 13.64 7. Hubert Grossard ([42 42 Belgium 42]) 13.65 8. Jonathan N'Senga ([43 43 Belgium 43]) 13.66 9. Johan Lisabeth ([44 44 Belgium 44]) 13.75 Women 's 5,000 metres 1. [45 45 Roberta Brunet 45] ([46 46 Italy 46]) 14 minutes 48.96 seconds 2. [47 47 Fernanda Ribeiro 47] ([48 48 Portugal 48]) 14:49.81 3. [49 49 Sally Barsosio 49] ([50 50 Kenya 50])

) 14:58.29 4. [51 51 Paula Radcliffe 51] ([52 52 Britain 52]) 14:59.70 5. Julia Vaquero ([53 53 Spain 53]) 15:04.94 6. Catherine McKiernan ([54 54 Ireland 54]) 15:07.57 7. Annette Peters ([55 55 U.S. 55]) 15:07.85 8. [56 56 Pauline Konga 56] ([57 57 Kenya 57]) 15:11.40 Men 's 100 metres 1. [58 58 Dennis Mitchell 58] ([59 59 U.S. 59]) 10.03 seconds 2. [60 60 Donovan Bailey 60] ([61 61 Canada 61]) 10.09 3. [62 62 Carl Lewis 62] ([63 63 U.S. 63]) 10.10 4. [64 64 Ato Boldon 64] (Trinidad) 10.12 5. [65 65 Linford Christie 65] ([66 66 Britain 66]) 10.14 6. [67 67 Davidson Ezinwa 67] ([68 68 Nigeria 68]) 10.15 7. [69 69 Jon Drummond 69] ([70 70 U.S. 70]) 10.16 8. [71 71 Bruny Surin 71] ([72 72 Canada 72]) 10.30 Men 's 400 metres hurdles 1. [73 73 Derrick Adkins 73] ([74 74 U.S. 74]) 47.93 seconds 2. [75 75 Samuel Matete 75] ([76 76 Zambia 76]) 47.99 3. [77 77 Rohan Robinson 77] ([78 78 Australia 78]) 48.86 4. [79 79 Torrance Zellner 79] ([80 80 U.S. 80]) 49.06 5. Jean-Paul Bruwier ([81 81 Belgium 81]) 49.24 6. Dusan Kovacs ([82 82 Hungary 82]) 49.31 7. [83 83 Calvin Davis 83] ([84 84 U.S. 84]) 49.49 8. [85 85 Laurent Ottoz 85] ([86 86 Italy 86]) 49.61 9. Marc Dollendorf ([87 87 Belgium 87]) 50.36 Women 's 100 metres 1. [88 88 Gail Devers 88] ([89 89 U.S. 89]) 10.84 seconds 2. [90 90 Gwen Torrence 90] ([91 91 U.S. 91]) 11.00 3. [92 92 Merlene Ottey 92] ([93 93 Jamaica 93]) 11.04 4. [94 94 Mary Onyali 94] ([95 95 Nigeria 95]) 11.09 5. [96 96 Chryste Gaines 96] ([97 97 U.S. 97]) 11.18 6. [98 98 Zhanna Pintusevich 98] ([99 99 Ukraine 99]) 11.27 7. [100 100 Irina Privalova 100] ([101 101 Russia 101]) 11.28 8. Natalia Voronova ([102 102 Russia 102]) 11.28 9. [103 103 Juliet Cuthbert 103] ([104 104 Jamaica 104]) 11.31 Women 's 1,500 metres 1. [105 105 Regina Jacobs 105] ([106 106 U.S. 106]) 4 minutes 01.77 seconds 2. Patricia Djate ([107 107 France 107]) 4:02.26 3. [108 108 Carla Sacramento 108] ([109 109 Portugal 109]) 4:02.67 4. [110 110 Yekaterina Podkopayeva 110] ([111 111 Russia 111]) 4:04.78 5. Margret Crowley ([112 112 Australia 112]) 4:05.00 6. [113 113 Leah Pells 113] ([114 114 Canada 114]) 4:05.64 7. [115 115 Sarah Thorsett 115] ([116 116 U.S. 116]) 4:06.80 8. Sinead Delahunty ([117 117 Ireland 117]) 4:07.27 3,000 metres steeplechase 1. [118 118 Joseph Keter 118] ([119 119 Kenya 119]) 8 minutes 10.02 seconds 2. [120 120 Patrick Sang 120] ([121 121 Kenya 121]) 8:12.04 3. [122 122 Moses Kiptanui 122] ([123 123 Kenya 123]) 8:12.65 4. Gideon Chirchir ([124 124 Kenya 124]) 8:15.69 5. Richard Kosgei ([125 125 Kenya 125]) 8:16.80 6. Larbi El Khattabi ([126 126 Morocco 126]) 8:17.29 7. Eliud Barngetuny ([127 127 Kenya 127]) 8:17.66 8. [128 128 Bernard Barmasai 128] ([129 129 Kenya 129]) 8:17.94 Men 's 400 metres 1. [130 130 Michael Johnson 130] ([131 131 U.S. 131]) 44.29 seconds 2. [132 132 Derek Mills 132] ([133 133 U.S. 133]) 44.78 3. [134 134 Anthuan Maybank 134] ([135 135 U.S. 135]) 44.92 4. [136 136 Davis Kamoja 136] ([137 137 Uganda 137]) 44.96 5. [138 138 Jamie Baulch 138] ([139 139 Britain 139]) 45.08 6. [140 140 Sunday Ba

da 140] ([141 141 Nigeria 141]) 45.21 7. [142 142 Samson Kitur 142] ([143 143 Kenya 143]) 45.34 8. [144 144 Mark Richardson 144] ([145 145 Britain 145]) 45.67 9. [146 146 Jason Rousier 146] ([147 147 U.S. 147]) 46.11 Men 's 200 metres 1. [148 148 Frankie Fredericks 148] ([149 149 Namibia 149]) 19.92 seconds 2. [150 150 Ato Boldon 150] (Trinidad) 19.99 3. [151 151 Jeff Williams 151] ([152 152 U.S. 152]) 20.21 4. [153 153 Jon Drummond 153] ([154 154 U.S. 154]) 20.42 5. [155 155 Patrick Stevens 155] ([156 156 Belgium 156]) 20.42 6. [157 157 Michael Marsh 157] ([158 158 U.S. 158]) 20.43 7. [159 159 Ivan Garcia 159] ([160 160 Cuba 160]) 20.45 8. Eric Wymeersch ([161 161 Belgium 161]) 20.84 9. [162 162 La mont Smith 162] ([163 163 U.S. 163]) 21.08 Women 's 1,000 metres 1. [164 164 Svetlana Masterkova 164] ([165 165 Russia 165]) 2 minutes 28.98 seconds (world record) 2. [166 166 Maria Mutola 166] ([167 167 Mozambique 167]) 2:29.66 3. [168 168 Malgorzata Rydz 168] ([169 169 Poland 169]) 2:39.00 4. Anja Smolders ([170 170 Belgium 170]) 2:43.06 5. Veerle De Jaeghere ([171 171 Belgium 171]) 2:43.18 6. Eleonora Berlanda ([172 172 Italy 172]) 2:43.44 7. Anneke Matthijs ([173 173 Belgium 173]) 2:43.82 8. Jacqueline Martin ([174 174 Spain 174]) 2:44.22 Women 's 200 metres 1. [175 175 Mary On yali 175] ([176 176 Nigeria 176]) 22.42 seconds 2. [177 177 Inger Miller 177] ([178 178 U.S. 178]) 22.66 3. [179 179 Irina Privalova 179] ([180 180 Russia 180]) 22.68 4. Natalia Voronova ([181 181 Russia 181]) 22.73 5. [182 182 Marina Trandenkova 182] ([183 183 Russia 183]) 22.84 6. [184 184 Chandra Sturrup 184] ([185 185 Bahamas 185]) 22.85 7. Zundra Feagin ([186 186 U.S. 186]) 23.18 8. [187 187 Galina Malchugina 187] ([188 188 Russia 188]) 23.25 Women 's 400 metres 1. [189 189 Cathy Freeman 189] ([190 190 Australia 190]) 49.48 seconds 2. [191 191 Marie-Jose Perec 191] ([192 192 France 192]) 49.72 3. [193 193 Falilat Ogunkoya 193] ([194 194 Nigeria 194]) 49.97 4. [195 195 Pauline Davis 195] ([196 196 Bahamas 196]) 50.14 5. Fatima Yussuf ([197 197 Nigeria 197]) 50.14 6. Maicel Malone ([198 198 U.S. 198]) 50.51 7. Hana Benesova ([199 199 Czech Republic 199]) 51.71 8. Ann Mercken ([200 200 Belgium 200]) 53.55 Men 's 3,000 metres 1. [201 201 Daniel Komen 201] ([202 202 Kenya 202]) 7 minutes 25.87 seconds 2. [203 203 Khalid Boulami 203] ([204 204 Morocco 204]) 7:31.65 3. [205 205 Bob Kennedy 205] ([206 206 U.S. 206]) 7:31.69 4. El Hassane Lahssini ([207 207 Morocco 207]) 7:32.44 5. [208 208 Thomas Nyariki 208] ([209 209 Kenya 209]) 7:35.56 6. [210 210 Noureddine Morceli 210] ([211 211 Algeria 211]) 7:36.81 7. Fita Bayesa ([212 212 Ethiopia 212]) 7:38.09 8. Martin Keino ([213 213 Kenya 213]) 7:38.88 Men 's discus 1. [214 214 Lars Riedel 214] ([215 215 Germany 215]) 66.74 metres 2. [216 216 Anthony Washington 216] ([217 217 U.S. 217]) 66.72 3. [218 218 Vladimir Dubrovshchik 218] ([219 219 Belarus 219]) 64.02 4. Virgilius Alekna ([220 220 Lithuania 220]) 63.62 5. [221 221 Juergen Schult 221] ([222 222 Germany 222]) 63.48 6. Vassiliy Kaptyukh ([223 223 Ukraine 223]) 63.48

3 223 Belarus 223]) 61.80 7. Vaclavas Kidikas ([224 224 Lithuania 224]) 60.92 8. [225 225 Michael Mollenbeck 225] ([226 226 Germany 226]) 59.24 Men 's triple jump 1. [227 227 Jonatan Edwards 227] ([228 228 Britain 228]) 17.50 metres 2. [229 229 Yoelvis Quesada 229] ([230 230 Cuba 230]) 17.29 3. [231 231 Brian Wellman 231] ([232 232 Bermuda 232]) 17.05 4. [233 233 Kenny Harrison 233] ([234 234 U.S. 234]) 16.97 5. Gennadi Markov ([235 235 Russia 235]) 16.66 6. [236 236 Francis Agyepong 236] ([237 237 Britain 237]) 16.63 7. [238 238 Rogel Nachum 238] ([239 239 Israel 239]) 16.36 8. [240 240 Sigurd Njerve 240] ([241 241 Norway 241]) 16.35 Men 's 1,500 metres 1. [242 242 Hicham El Guerrouj 242] ([243 243 Morocco 243]) three minutes 29.05 seconds 2. [244 244 Isaac Viciosa 244] ([245 245 Spain 245]) 3:33.00 3. [246 246 William Tanui 246] ([247 247 Kenya 247]) 3:33.36 4. Elijah Maru ([248 248 Kenya 248]) 3:33.64 5. Marcus O'Sullivan ([249 249 Ireland 249]) 3:33.77 6. [250 250 John Mayock 250] ([251 251 Britain 251]) 3:33.94 7. [252 252 Laban Rotich 252] ([253 253 Kenya 253]) 3:34.12 8. Christophe Impens ([254 254 Belgium 254]) 3:34.13 Women 's high jump 1. [255 255 Stefka Kostadinova 255] ([256 256 Bulgaria 256]) 2.03 metres 2. [257 257 Inga Babanova 257] ([258 258 Ukraine 258]) 2.03 3. [259 259 Alina Astafei 259] ([260 260 Germany 260]) 1.97 4. [261 261 Tatyana Moktova 261] ([262 262 Russia 262]) 1.94 5. [263 263 Svetlana Zalevskaya 263] ([264 264 Kazakhstan 264]) 1.91 6. [265 265 Yelena Gulyayeva 265] ([266 266 Russia 266]) 1.88 7. Hanna Haugland ([267 267 Norway 267]) 1.88 8 equal . Olga Boshova ([268 268 Moldova 268]) 1.85 8 equal . [269 269 Nele Zilins kiene 269] ([270 270 Lithuania 270]) 1.85 Men 's 10,000 metres 1. [271 271 Salah Hissou 271] ([272 272 Morocco 272]) 26 minutes 38.08 seconds (world record) 2. [273 273 Paul Tergat 273] ([274 274 Kenya 274]) 26:54.41 3. [275 275 Paul Koech 275] ([276 276 Kenya 276]) 26:56.78 4. William Kiptum ([277 277 Kenya 277]) 27:18.84 5. [278 278 Aloys Nzigama 278] ([279 279 Burundi 279]) 27:25.13 6. [280 280 Mathias Ntawulikura 280] ([281 281 Rwanda 281]) 27:25.48 7. [282 282 Abel Anton 282] ([283 283 Spain 283]) 28:18.44 8. [284 284 Kamiel Maase 284] ([285 285 Netherlands 285]) 28.29.42 9. Worku Bekila ([286 286 Ethiopia 286]) 28.42.23 10. Robert Steffko ([287 287 Slovakia 287]) 28:42.26

- 1: 3708 Brussels
- 2: 2738276 Ilke_Wyludda
- 3: 11867 Germany
- 4: 2731709 Ellina_Zvereva
- 5: 3457 Belarus
- 6: 3196764 Franka_Dietzschen
- 7: 11867 Germany
- 8: 3909612 Natalya_Sadova
- 9: 25391 Russia
- 10: 8481614 Mette_Bergmann
- 11: 21241 Norway
- 12: 3026832 Nicoleta_Grasu

13: 25445 Romania
14: 3518703 Olga_Chernyavskaya
15: 25391 Russia
16: 2913184 Iryna_Yatchenko
17: 3457 Belarus
18: 30873921 Ludmila_Engquist
19: 5058739 Sweden
20: 6829072 Michelle_Freeman
21: 15660 Jamaica
22: 3020050 Aliuska_López
23: 5042481 Cuba
24: 15660 Jamaica
25: 4273494 Brigit_Bukovec
26: 26830 Slovakia
27: 25391 Russia
28: 26748 Switzerland
29: 6827139 Patricia_Girard
30: 5843419 France
31: 3434750 United_States
32: 802066 Allen_Johnson
33: 3434750 United_States
34: 789006 Colin_Jackson
35: 31717 United_Kingdom
36: 11849304 Emilio_Valle
37: 5042481 Cuba
38: 21957117 Sven_Pieters
39: 3343 Belgium
40: 3434750 United_States
41: 3343 Belgium
42: 3343 Belgium
43: 3343 Belgium
44: 3343 Belgium
45: 2689445 Roberta_Brunet
46: 14532 Italy
47: 1787811 Fernanda_Ribeiro
48: 23033 Portugal
49: 2531356 Sally_Barsosio
50: 188171 Kenya
51: 228240 Paula_Radcliffe
52: 31717 United_Kingdom
53: 26667 Spain
54: 14560 Republic_of_Ireland
55: 3434750 United_States
56: 2456590 Pauline_Konga
57: 188171 Kenya
58: 1148663 Dennis_Mitchell
59: 3434750 United_States
60: 67116 Donovan_Bailey
61: 5042916 Canada
62: 67127 Carl_Lewis
63: 3434750 United_States
64: 96607 Ato_Boldon

65: 67118 Linford_Christie
66: 31717 United_Kingdom
67: 2732799 Davidson_Ezinwa
68: 21383 Nigeria
69: 1157765 Jon_Drummond
70: 3434750 United_States
71: 940384 Bruny_Surin
72: 5042916 Canada
73: 993078 Derrick_Adkins
74: 3434750 United_States
75: 3020977 Samuel_Matete
76: 34415 Zambia
77: 12295262 Rohan_Robinson
78: 4689264 Australia
79: 2530948 Torrance_Zellner
80: 3434750 United_States
81: 3343 Belgium
82: 13275 Hungary
83: 7457007 Calvin_Davis
84: 3434750 United_States
85: 11880301 Laurent_Ottoz
86: 14532 Italy
87: 3343 Belgium
88: 72654 Gail_Devers
89: 3434750 United_States
90: 882492 Gwen_Torrence
91: 3434750 United_States
92: 558060 Merlene_Ottey
93: 15660 Jamaica
94: 2722641 Mary_Onyali-Omagbemi
95: 21383 Nigeria
96: 6812537 Chryste_Gaines
97: 3434750 United_States
98: 30874167 Zhanna_Pintusevich-Block
99: 31750 Ukraine
100: 49242 Irina_Privalova
101: 25391 Russia
102: 25391 Russia
103: 6756205 Juliet_Cuthbert
104: 15660 Jamaica
105: 6192340 Regina_Jacobs
106: 3434750 United_States
107: 5843419 France
108: 2373750 Carla_Sacramento
109: 23033 Portugal
110: 3857148 Yekaterina_Podkopayeva
111: 25391 Russia
112: 4689264 Australia
113: 13926383 Leah_Pells
114: 5042916 Canada
115: 21758160 Sarah_Thorsett
116: 3434750 United_States

117: 14560 Republic_of_Ireland
118: 1036695 Joseph_Keter
119: 188171 Kenya
120: 4603285 Patrick_Sang
121: 188171 Kenya
122: 857479 Moses_Kiptanui
123: 188171 Kenya
124: 188171 Kenya
125: 188171 Kenya
126: 19291 Morocco
127: 188171 Kenya
128: 9577738 Bernard_Barmasai
129: 188171 Kenya
130: 653621 Michael_Johnson_(sprinter)
131: 3434750 United_States
132: 1676095 Derek_Mills
133: 3434750 United_States
134: 1676098 Anthuan_Maybank
135: 3434750 United_States
136: 2434413 Davis_Kamoga
137: 31816 Uganda
138: 773005 Jamie_Baulch
139: 31717 United_Kingdom
140: 2738815 Sunday_Bada
141: 21383 Nigeria
142: 289809 Samson_Kitur
143: 188171 Kenya
144: 7652893 Mark_Richardson_(athlete)
145: 31717 United_Kingdom
146: 13121324 Jason_Rouser
147: 3434750 United_States
148: 71410 Frankie_Fredericks
149: 21292 Namibia
150: 96607 Ato_Boldon
151: 27446348 Jeff_Williams_(athlete)
152: 3434750 United_States
153: 1157765 Jon_Drummond
154: 3434750 United_States
155: 13974630 Patrick_Stevens
156: 3343 Belgium
157: 681338 Michael_Marsh_(athlete)
158: 3434750 United_States
159: 2688112 Iván_García_(athlete)
160: 5042481 Cuba
161: 3343 Belgium
162: 1676088 LaMont_Smith
163: 3434750 United_States
164: 613736 Svetlana_Masterkova
165: 25391 Russia
166: 66748 Maria_Mutola
167: 19301 Mozambique
168: 21285670 Małgorzata_Rydz

169: 22936 Poland
170: 3343 Belgium
171: 3343 Belgium
172: 14532 Italy
173: 3343 Belgium
174: 26667 Spain
175: 2722641 Mary_Onyali-Omagbemi
176: 21383 Nigeria
177: 1031484 Inger_Miller
178: 3434750 United_States
179: 49242 Irina_Privalova
180: 25391 Russia
181: 25391 Russia
182: 6812531 Marina_Trandenkova
183: 25391 Russia
184: 2511007 Chandra_Sturrup
185: 3451 The_Bahamas
186: 3434750 United_States
187: 3524541 Galina_Malchugina
188: 25391 Russia
189: 66002 Cathy_Freeman
190: 4689264 Australia
191: 66161 Marie-José_Pérec
192: 5843419 France
193: 2489403 Falilat_Ogunkoya
194: 21383 Nigeria
195: 3670385 Pauline_Davis-Thompson
196: 3451 The_Bahamas
197: 21383 Nigeria
198: 3434750 United_States
199: 5321 Czech_Republic
200: 3343 Belgium
201: 673604 Daniel_Komen
202: 188171 Kenya
203: 3517130 Khalid_Boulami
204: 19291 Morocco
205: 3058737 Bob_Kennedy_(athlete)
206: 3434750 United_States
207: 19291 Morocco
208: 2736339 Tom_Nyariki
209: 188171 Kenya
210: 492499 Noureddine_Morceli
211: 358 Algeria
212: 187749 Ethiopia
213: 188171 Kenya
214: 3908095 Lars_Riedel
215: 11867 Germany
216: 6781457 Anthony_Washington
217: 3434750 United_States
218: 3310705 Vladimir_Dubrovshchik
219: 3457 Belarus
220: 17675 Lithuania

221: 3918291 Jürgen_Schult
222: 11867 Germany
223: 3457 Belarus
224: 17675 Lithuania
225: 3026211 Michael_Möllenbeck
226: 11867 Germany
227: 52056 Jonathan_Edwards_(athlete)
228: 31717 United_Kingdom
229: 3241060 Yoelbi_Quesada
230: 5042481 Cuba
231: 2690829 Brian_Wellman
232: 3460 Bermuda
233: 2645623 Kenny_Harrison
234: 3434750 United_States
235: 25391 Russia
236: 5912006 Francis_Agyepong
237: 31717 United_Kingdom
238: 9788550 Rogel_Nachum
239: 9282173 Israel
240: 8485751 Sigurd_Njerve
241: 21241 Norway
242: 492488 Hicham_El_Guerrouj
243: 19291 Morocco
244: 8903845 Isaac_Viciosa
245: 26667 Spain
246: 815988 William_Tanui
247: 188171 Kenya
248: 188171 Kenya
249: 14560 Republic_of_Ireland
250: 13182495 John_Mayock
251: 31717 United_Kingdom
252: 6261028 Laban_Rotich
253: 188171 Kenya
254: 3343 Belgium
255: 826515 Stefka_Kostadinova
256: 3415 Bulgaria
257: 6299154 Inha_Babakova
258: 31750 Ukraine
259: 3905854 Alina_Astafei
260: 11867 Germany
261: 8278156 Tatyana_Babashkina
262: 25391 Russia
263: 8277997 Svetlana_Zalevskaya
264: 16642 Kazakhstan
265: 8278273 Yelena_Gulyayeva
266: 25391 Russia
267: 21241 Norway
268: 19260 Moldova
269: 10906630 Nelė_Žilinskienė
270: 17675 Lithuania
271: 3580285 Salah_Hissou
272: 19291 Morocco

273: 597066 Paul_Tergat
274: 188171 Kenya
275: 7080613 Paul_Koech
276: 188171 Kenya
277: 188171 Kenya
278: 7190467 Aloës_Nizigama
279: 21490998 Burundi
280: 7190531 Mathias_Ntawulikura
281: 25645 Rwanda
282: 3552171 Abel_Antón
283: 26667 Spain
284: 8380236 Kamiel_Maase
285: 21148 Netherlands
286: 187749 Ethiopia
287: 26830 Slovakia

1: 7885102 Steve_Brown_(athlete)
[winner: 7885102 Steve_Brown_(athlete), pem_pos=13,
final_scores=0.015175685, final_local_scores=0.10283664,
pem=-2.6390574, similarity=21.510168]
gt: 7885102 Steve_Brown_(athlete), pem_pos=13,
final_scores=0.015175685, final_local_scores=0.10283664,
pem=-2.6390574, similarity=21.510168]

1: 3708 Brussels
[pem_pos=0, final_scores=0.5474059,
final_local_scores=0.5696541, pem=-0.025440974,
similarity=21.22768]

2: 2738276 Ilke_Wyludda
[pem_pos=0, final_scores=0.48606938,
final_local_scores=0.66181284, pem=0.0, similarity=23.839529]

3: 11867 Germany
[pem_pos=0, final_scores=0.8350087,
final_local_scores=0.8526873, pem=-0.15529114,
similarity=30.3801]

4: 2731709 Ellina_Zvereva
[pem_pos=0, final_scores=0.6842722,
final_local_scores=0.85927415, pem=0.0, similarity=29.732168]

5: 3457 Belarus
[pem_pos=0, final_scores=0.6252734,
final_local_scores=0.6241517, pem=-0.4220667,
similarity=25.010908]

6: 3196764 Franka_Dietzsch
[pem_pos=0, final_scores=0.85292906,
final_local_scores=1.0161717, pem=0.0, similarity=34.414307]

7: 11867 Germany
[pem_pos=0, final_scores=0.8438836,
final_local_scores=0.8623819, pem=-0.15529114,
similarity=30.669407]

8: 3909612 Natalya_Sadova
[pem_pos=0, final_scores=0.62667215,
final_local_scores=0.78850996, pem=0.0, similarity=27.620422]

9: 25391 Russia
[pem_pos=0, final_scores=0.82783777,

```
final_local_scores=0.85389394, pem=-0.20299491,
similarity=30.67553]
10: 8481614 Mette_Bergmann
    [pem_pos=0, final_scores=0.4933812,
final_local_scores=0.68445456, pem=0.0, similarity=24.515202]
11: 21241 Norway
    [pem_pos=0, final_scores=0.77990896,
final_local_scores=0.82163596, pem=-0.22668563,
similarity=29.84172]
12: 3026832 Nicoleta_Grasu
    [pem_pos=0, final_scores=0.90739137,
final_local_scores=1.0719976, pem=0.0, similarity=36.08026]
13: 25445 Romania
    [pem_pos=0, final_scores=0.70653564,
final_local_scores=0.7182409, pem=-0.21241479,
similarity=26.678598]
14: 3518703 Olga_Chernyavskaya
    [pem_pos=0, final_scores=0.6403526,
final_local_scores=0.7907993, pem=0.0, similarity=27.688742]
15: 25391 Russia
    [pem_pos=0, final_scores=0.82418233,
final_local_scores=0.8499009, pem=-0.20299491,
similarity=30.556368]
16: 2913184 Iryna_Yatchenko
    [pem_pos=0, final_scores=0.7738078,
final_local_scores=0.91134673, pem=0.0, similarity=31.286118]
17: 3457 Belarus
    [pem_pos=0, final_scores=0.58449394,
final_local_scores=0.5796063, pem=-0.4220667,
similarity=23.681585]
18: 30873921 Ludmila_Engquist
    [pem_pos=0, final_scores=0.60348755,
final_local_scores=0.74202657, pem=0.0, similarity=26.233265]
19: 5058739 Sweden
    [pem_pos=0, final_scores=0.7666411,
final_local_scores=0.78335226, pem=-0.2135508,
similarity=28.627829]
20: 6829072 Michelle_Freeman
    [pem_pos=0, final_scores=0.6342151,
final_local_scores=0.7911869, pem=0.0, similarity=27.70031]
21: 15660 Jamaica
    [pem_pos=0, final_scores=0.5138704,
final_local_scores=0.5026275, pem=-0.33295843,
similarity=20.8998]
22: 3020050 Aliuska_López
    [pem_pos=0, final_scores=0.5694169,
final_local_scores=0.72477186, pem=0.0, similarity=25.718351]
23: 5042481 Cuba
    [pem_pos=0, final_scores=0.5479903,
final_local_scores=0.54389185, pem=-0.23907764,
similarity=21.620672]
24: 15660 Jamaica
```

[pem_pos=0, final_scores=0.49876553,
final_local_scores=0.48612773, pem=-0.33295843,
similarity=20.407413]
25: 4273494 Brigitा_Bukovec
[pem_pos=0, final_scores=0.751255,
final_local_scores=0.929303, pem=0.0, similarity=31.821966]
26: 26830 Slovakia
[pem_pos=0, final_scores=0.668178,
final_local_scores=0.67863435, pem=-0.36602598,
similarity=26.332022]
27: 25391 Russia
[pem_pos=0, final_scores=0.81629056,
final_local_scores=0.8412804, pem=-0.20299491,
similarity=30.299116]
28: 26748 Switzerland
[pem_pos=0, final_scores=0.6671379,
final_local_scores=0.66997063, pem=-0.18598008,
similarity=25.09436]
29: 6827139 Patricia_Girard
[pem_pos=0, final_scores=0.5448535,
final_local_scores=0.6785737, pem=0.0, similarity=24.339706]
30: 5843419 France
[pem_pos=0, final_scores=0.76236564,
final_local_scores=0.7716152, pem=-0.16240834,
similarity=27.99945]
31: 3434750 United_States
[pem_pos=0, final_scores=0.51322293,
final_local_scores=0.48987728, pem=-0.14075686,
similarity=19.474087]
32: 802066 Allen_Johnson
[pem_pos=0, final_scores=0.5596806,
final_local_scores=0.6542439, pem=0.0, similarity=23.613655]
33: 3434750 United_States
[pem_pos=0, final_scores=0.50563353,
final_local_scores=0.481587, pem=-0.14075686,
similarity=19.226686]
34: 789006 Colin_Jackson
[pem_pos=0, final_scores=0.6263377,
final_local_scores=0.69529724, pem=0.0, similarity=24.83877]
35: 31717 United_Kingdom
[pem_pos=0, final_scores=0.59699255,
final_local_scores=0.58778125, pem=-0.61114657,
similarity=24.953789]
36: 11849304 Emilio_Valle
[pem_pos=0, final_scores=0.54071075,
final_local_scores=0.7150549, pem=0.0, similarity=25.42838]
37: 5042481 Cuba
[pem_pos=0, final_scores=0.54178077,
final_local_scores=0.53710884, pem=-0.23907764,
similarity=21.418253]
38: 21957117 Sven_Pieters
[pem_pos=0, final_scores=0.43678284,

```
final_local_scores=0.5724065, pem=0.0, similarity=21.171463]
39: 3343 Belgium
    [pem_pos=0, final_scores=0.73209625,
final_local_scores=0.73459435, pem=-0.24180195,
similarity=27.326431]
40: 3434750 United_States
    [pem_pos=0, final_scores=0.5037344,
final_local_scores=0.4795125, pem=-0.14075686,
similarity=19.164782]
41: 3343 Belgium
    [pem_pos=0, final_scores=0.72862405,
final_local_scores=0.73080146, pem=-0.24180195,
similarity=27.213243]
42: 3343 Belgium
    [pem_pos=0, final_scores=0.69112176,
final_local_scores=0.6898359, pem=-0.24180195,
similarity=25.99075]
43: 3343 Belgium
    [pem_pos=0, final_scores=0.70851725,
final_local_scores=0.7088378, pem=-0.24180195,
similarity=26.557804]
44: 3343 Belgium
    [pem_pos=0, final_scores=0.7126463,
final_local_scores=0.71334815, pem=-0.24180195,
similarity=26.692402]
45: 2689445 Roberta_Brunet
    [pem_pos=0, final_scores=0.5496691,
final_local_scores=0.68808144, pem=0.0, similarity=24.623436]
46: 14532 Italy
    [pem_pos=0, final_scores=0.79442924,
final_local_scores=0.8240204, pem=-0.20612453,
similarity=29.801064]
47: 1787811 Fernanda_Ribeiro
    [pem_pos=0, final_scores=0.4712016,
final_local_scores=0.5702807, pem=0.0, similarity=21.108027]
48: 23033 Portugal
    [pem_pos=0, final_scores=0.5873236,
final_local_scores=0.5925283, pem=-0.30942732,
similarity=23.454653]
49: 2531356 Sally_Barsosio
    [pem_pos=0, final_scores=0.6855811,
final_local_scores=0.83120453, pem=0.0, similarity=28.894514]
50: 188171 Kenya
    [pem_pos=0, final_scores=0.6840735,
final_local_scores=0.7081201, pem=-0.27389407,
similarity=26.710907]
51: 228240 Paula_Radcliffe
    [pem_pos=0, final_scores=0.59095687,
final_local_scores=0.66432416, pem=0.0, similarity=23.91447]
52: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.584693,
final_local_scores=0.5743459, pem=-0.61114657,
```

```
similarity=24.552849]
53: 26667 Spain
    [pem_pos=0, final_scores=0.62488824,
final_local_scores=0.6213574, pem=-0.18857051,
similarity=23.657732]
54: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.5646518,
final_local_scores=0.5970114, pem=-1.5040284,
similarity=30.084866]
55: 3434750 United_States
    [pem_pos=0, final_scores=0.48889035,
final_local_scores=0.46329767, pem=-0.14075686,
similarity=18.680897]
56: 2456590 Pauline_Konga
    [pem_pos=0, final_scores=0.49485773,
final_local_scores=0.66401005, pem=0.0, similarity=23.905098]
57: 188171 Kenya
    [pem_pos=0, final_scores=0.6250172,
final_local_scores=0.64361006, pem=-0.27389407,
similarity=24.785799]
58: 1148663 Dennis_Mitchell
    [pem_pos=0, final_scores=0.556444,
final_local_scores=0.65808326, pem=0.0, similarity=23.72823]
59: 3434750 United_States
    [pem_pos=0, final_scores=0.48465842,
final_local_scores=0.45867485, pem=-0.14075686,
similarity=18.542942]
60: 67116 Donovan_Bailey
    [pem_pos=0, final_scores=0.73312587,
final_local_scores=0.84603655, pem=0.0, similarity=29.337132]
61: 5042916 Canada
    [pem_pos=0, final_scores=0.7819589,
final_local_scores=0.7914228, pem=-0.11662793,
similarity=28.341587]
62: 67127 Carl_Lewis
    [pem_pos=0, final_scores=0.72746444,
final_local_scores=0.81206477, pem=0.0, similarity=28.323345]
63: 3434750 United_States
    [pem_pos=0, final_scores=0.5219195,
final_local_scores=0.49937695, pem=-0.14075686,
similarity=19.757576]
64: 96607 Ato_Boldon
    [pem_pos=0, final_scores=0.78623825,
final_local_scores=0.87779206, pem=0.0, similarity=30.28478]
65: 67118 Linford_Christie
    [pem_pos=0, final_scores=0.929764,
final_local_scores=1.0460777, pem=0.0, similarity=35.30676]
66: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5840593,
final_local_scores=0.57365364, pem=-0.61114657,
similarity=24.53219]
67: 2732799 Davidson_Ezinwa
```

```
[pem_pos=0, final_scores=0.72255355,
final_local_scores=0.88128656, pem=0.0, similarity=30.389063]
68: 21383 Nigeria
    [pem_pos=0, final_scores=0.60081536,
final_local_scores=0.6167973, pem=-0.31455088,
similarity=24.206753]
69: 1157765 Jon_Drummond
    [pem_pos=0, final_scores=0.5184403,
final_local_scores=0.6363977, pem=0.0, similarity=23.08109]
70: 3434750 United_States
    [pem_pos=0, final_scores=0.48857892,
final_local_scores=0.46295744, pem=-0.14075686,
similarity=18.670746]
71: 940384 Bruny_Surin
    [pem_pos=0, final_scores=0.65716684,
final_local_scores=0.7919055, pem=0.0, similarity=27.721754]
72: 5042916 Canada
    [pem_pos=0, final_scores=0.7517852,
final_local_scores=0.75846267, pem=-0.11662793,
similarity=27.357992]
73: 993078 Derrick_Adkins
    [pem_pos=0, final_scores=0.561137,
final_local_scores=0.6845945, pem=0.0, similarity=24.519379]
74: 3434750 United_States
    [pem_pos=0, final_scores=0.5035594,
final_local_scores=0.4793213, pem=-0.14075686,
similarity=19.159073]
75: 3020977 Samuel_Matete
    [pem_pos=0, final_scores=0.7630301,
final_local_scores=0.89837056, pem=0.0, similarity=30.898884]
76: 34415 Zambia
    [pem_pos=0, final_scores=0.50771517,
final_local_scores=0.5367426, pem=-0.32073748,
similarity=21.851402]
77: 12295262 Rohan_Robinson
    [pem_pos=0, final_scores=0.6347541,
final_local_scores=0.7759728, pem=0.0, similarity=27.246288]
78: 4689264 Australia
    [pem_pos=0, final_scores=0.6960632,
final_local_scores=0.7044846, pem=-0.17168477,
similarity=26.046587]
79: 2530948 Torrance_Zellner
    [pem_pos=0, final_scores=0.549037,
final_local_scores=0.6948265, pem=0.0, similarity=24.82472]
80: 3434750 United_States
    [pem_pos=0, final_scores=0.5099419,
final_local_scores=0.48629326, pem=-0.14075686,
similarity=19.367132]
81: 3343 Belgium
    [pem_pos=0, final_scores=0.7417447,
final_local_scores=0.7451338, pem=-0.24180195,
similarity=27.64095]
```

82: 13275 Hungary
[pem_pos=0, final_scores=0.6973432,
final_local_scores=0.7055806, pem=-0.32577527,
similarity=26.917263]
83: 7457007 Calvin_Davis
[pem_pos=0, final_scores=0.35324854,
final_local_scores=0.49416763, pem=0.0, similarity=18.83666]
84: 3434750 United_States
[pem_pos=0, final_scores=0.48868054,
final_local_scores=0.46306843, pem=-0.14075686,
similarity=18.674057]
85: 11880301 Laurent_Ottoz
[pem_pos=0, final_scores=0.64317673,
final_local_scores=0.78802425, pem=0.0, similarity=27.605928]
86: 14532 Italy
[pem_pos=0, final_scores=0.73410505,
final_local_scores=0.7581254, pem=-0.20612453,
similarity=27.834625]
87: 3343 Belgium
[pem_pos=0, final_scores=0.726862,
final_local_scores=0.72887665, pem=-0.24180195,
similarity=27.155804]
88: 72654 Gail_Devers
[pem_pos=0, final_scores=0.7788206,
final_local_scores=0.90204865, pem=0.0, similarity=31.008644]
89: 3434750 United_States
[pem_pos=0, final_scores=0.5012905,
final_local_scores=0.47684294, pem=-0.14075686,
similarity=19.085115]
90: 882492 Gwen_Torrence
[pem_pos=0, final_scores=0.9456057,
final_local_scores=1.093163, pem=0.0, similarity=36.71188]
91: 3434750 United_States
[pem_pos=0, final_scores=0.5345046,
final_local_scores=0.5131243, pem=-0.14075686,
similarity=20.167824]
92: 558060 Merlene_Ottey
[pem_pos=0, final_scores=1.0239732,
final_local_scores=1.1434569, pem=0.0, similarity=38.212746]
93: 15660 Jamaica
[pem_pos=0, final_scores=0.52571064,
final_local_scores=0.51556116, pem=-0.33295843,
similarity=21.285767]
94: 2722641 Mary_Onyali-Omagbemi
[pem_pos=0, final_scores=0.7983522,
final_local_scores=0.94154435, pem=0.0, similarity=32.187275]
95: 21383 Nigeria
[pem_pos=0, final_scores=0.6416009,
final_local_scores=0.66134936, pem=-0.31455088,
similarity=25.536274]
96: 6812537 Chryste_Gaines
[pem_pos=0, final_scores=0.60147864,

```
final_local_scores=0.74498695, pem=0.0, similarity=26.32161]
97: 3434750 United_States
    [pem_pos=0, final_scores=0.5194922,
final_local_scores=0.4967255, pem=-0.14075686,
similarity=19.678452]
98: 30874167 Zhanna_Pintusevich-Block
    [pem_pos=0, final_scores=0.6189506,
final_local_scores=0.74451286, pem=0.0, similarity=26.30746]
99: 31750 Ukraine
    [pem_pos=0, final_scores=0.65168023,
final_local_scores=0.65162516, pem=-0.35407752,
similarity=25.461035]
100: 49242 Irina_Privalova
    [pem_pos=0, final_scores=0.79627985,
final_local_scores=0.9513189, pem=0.0, similarity=32.478966]
101: 25391 Russia
    [pem_pos=0, final_scores=0.8078701,
final_local_scores=0.83208233, pem=-0.20299491,
similarity=30.024626]
102: 25391 Russia
    [pem_pos=0, final_scores=0.84868854,
final_local_scores=0.8766703, pem=-0.20299491,
similarity=31.355219]
103: 6756205 Juliet_Cuthbert
    [pem_pos=0, final_scores=0.7949118,
final_local_scores=0.94404703, pem=0.0, similarity=32.26196]
104: 15660 Jamaica
    [pem_pos=0, final_scores=0.45720547,
final_local_scores=0.44072962, pem=-0.33295843,
similarity=19.052643]
105: 6192340 Regina_Jacobs
    [pem_pos=0, final_scores=0.540217,
final_local_scores=0.65579927, pem=0.0, similarity=23.66007]
106: 3434750 United_States
    [pem_pos=0, final_scores=0.46423537,
final_local_scores=0.43636578, pem=-0.14075686,
similarity=17.877197]
107: 5843419 France
    [pem_pos=0, final_scores=0.74317724,
final_local_scores=0.75065476, pem=-0.16240834,
similarity=27.37395]
108: 2373750 Carla_Sacramento
    [pem_pos=0, final_scores=0.43698883,
final_local_scores=0.5660802, pem=0.0, similarity=20.982676]
109: 23033 Portugal
    [pem_pos=0, final_scores=0.6447659,
final_local_scores=0.6552753, pem=-0.30942732,
similarity=25.327148]
110: 3857148 Yekaterina_Podkopayeva
    [pem_pos=0, final_scores=0.56753904,
final_local_scores=0.69566876, pem=0.0, similarity=24.849855]
111: 25391 Russia
```

```
[pem_pos=0, final_scores=0.8596111,
final_local_scores=0.8886016, pem=-0.20299491,
similarity=31.711273]
112: 4689264 Australia
    [pem_pos=0, final_scores=0.72873026,
final_local_scores=0.74016845, pem=-0.17168477,
similarity=27.111465]
113: 13926383 Leah_Pells
    [pem_pos=0, final_scores=0.4820274,
final_local_scores=0.6068905, pem=0.0, similarity=22.200537]
114: 5042916 Canada
    [pem_pos=0, final_scores=0.74710304,
final_local_scores=0.75334805, pem=-0.11662793,
similarity=27.205362]
115: 21758160 Sarah_Thorsett
    [pem_pos=0, final_scores=0.5150777,
final_local_scores=0.63369244, pem=0.0, similarity=23.000359]
116: 3434750 United_States
    [pem_pos=0, final_scores=0.51235205,
final_local_scores=0.488926, pem=-0.14075686,
similarity=19.445698]
117: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.56236714,
final_local_scores=0.59451574, pem=-1.5040284,
similarity=30.010391]
118: 1036695 Joseph_Keter
    [pem_pos=0, final_scores=0.41381764,
final_local_scores=0.55359614, pem=0.0, similarity=20.610126]
119: 188171 Kenya
    [pem_pos=0, final_scores=0.63625556,
final_local_scores=0.65588635, pem=-0.27389407,
similarity=25.152147]
120: 4603285 Patrick_Sang
    [pem_pos=0, final_scores=0.5085626,
final_local_scores=0.62276024, pem=0.0, similarity=22.67412]
121: 188171 Kenya
    [pem_pos=0, final_scores=0.65364593,
final_local_scores=0.6748826, pem=-0.27389407,
similarity=25.719034]
122: 857479 Moses_Kiptanui
    [pem_pos=0, final_scores=0.75235945,
final_local_scores=0.859846, pem=0.0, similarity=29.749233]
123: 188171 Kenya
    [pem_pos=0, final_scores=0.59836096,
final_local_scores=0.61449224, pem=-0.27389407,
similarity=23.916864]
124: 188171 Kenya
    [pem_pos=0, final_scores=0.6391277,
final_local_scores=0.6590236, pem=-0.27389407,
similarity=25.24577]
125: 188171 Kenya
    [pem_pos=0, final_scores=0.63277227,
```

```
final_local_scores=0.6520813, pem=-0.27389407,
similarity=25.038599]
126: 19291 Morocco
    [pem_pos=0, final_scores=0.648863,
final_local_scores=0.6656631, pem=-0.29598162,
similarity=25.564022]
127: 188171 Kenya
    [pem_pos=0, final_scores=0.6406774,
final_local_scores=0.6607165, pem=-0.27389407,
similarity=25.29629]
128: 9577738 Bernard_Barmasai
    [pem_pos=0, final_scores=0.5892665,
final_local_scores=0.72035533, pem=0.0, similarity=25.586554]
129: 188171 Kenya
    [pem_pos=0, final_scores=0.6348409,
final_local_scores=0.65434104, pem=-0.27389407,
similarity=25.106033]
130: 653621 Michael_Johnson_(sprinter)
    [pem_pos=0, final_scores=0.5578851,
final_local_scores=0.63846976, pem=-0.36135122,
similarity=25.108007]
131: 3434750 United_States
    [pem_pos=0, final_scores=0.48366988,
final_local_scores=0.45759505, pem=-0.14075686,
similarity=18.510721]
132: 1676095 Derek_Mills
    [pem_pos=0, final_scores=0.4812448,
final_local_scores=0.61665785, pem=0.0, similarity=22.492014]
133: 3434750 United_States
    [pem_pos=0, final_scores=0.49156898,
final_local_scores=0.46622366, pem=-0.14075686,
similarity=18.768213]
134: 1676098 Anthuan_Maybank
    [pem_pos=0, final_scores=0.55767816,
final_local_scores=0.69186974, pem=0.0, similarity=24.736486]
135: 3434750 United_States
    [pem_pos=0, final_scores=0.5098161,
final_local_scores=0.4861558, pem=-0.14075686,
similarity=19.36303]
136: 2434413 Davis_Kamoga
    [pem_pos=0, final_scores=0.5621369,
final_local_scores=0.687465, pem=0.0, similarity=24.60504]
137: 31816 Uganda
    [pem_pos=0, final_scores=0.6008912,
final_local_scores=0.6265399, pem=-0.2971513,
similarity=24.402866]
138: 773005 Jamie_Baulch
    [pem_pos=0, final_scores=0.67565465,
final_local_scores=0.7567827, pem=0.0, similarity=26.673618]
139: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.59767276,
final_local_scores=0.5885243, pem=-0.61114657,
```

```
similarity=24.97596]
140: 2738815 Sunday_Bada
    [pem_pos=0, final_scores=0.6324623,
final_local_scores=0.743923, pem=0.0, similarity=26.28986]
141: 21383 Nigeria
    [pem_pos=0, final_scores=0.63329095,
final_local_scores=0.652272, pem=-0.31455088,
similarity=25.265387]
142: 289809 Samson_Kitur
    [pem_pos=0, final_scores=0.5942272,
final_local_scores=0.7211523, pem=0.0, similarity=25.610336]
143: 188171 Kenya
    [pem_pos=0, final_scores=0.64999825,
final_local_scores=0.67089814, pem=-0.27389407,
similarity=25.60013]
144: 7652893 Mark_Richardson_(athlete)
    [pem_pos=2, final_scores=0.48089308,
final_local_scores=0.59061545, pem=-1.0948459,
similarity=27.668798]
145: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5879312,
final_local_scores=0.57788306, pem=-0.61114657,
similarity=24.658407]
146: 13121324 Jason_Rouser
    [pem_pos=0, final_scores=0.5834946,
final_local_scores=0.72029203, pem=0.0, similarity=25.584663]
147: 3434750 United_States
    [pem_pos=0, final_scores=0.4624117,
final_local_scores=0.43437368, pem=-0.14075686,
similarity=17.817749]
148: 71410 Frankie_Fredericks
    [pem_pos=0, final_scores=0.752279,
final_local_scores=0.87179464, pem=0.0, similarity=30.105803]
149: 21292 Namibia
    [pem_pos=0, final_scores=0.47620845,
final_local_scores=0.49531662, pem=-0.36777115,
similarity=20.870945]
150: 96607 Ato_Boldon
    [pem_pos=0, final_scores=0.8429952,
final_local_scores=0.9397904, pem=0.0, similarity=32.134933]
151: 27446348 Jeff_Williams_(athlete)
    [pem_pos=5, final_scores=0.3990329,
final_local_scores=0.48926407, pem=-1.5652602,
similarity=27.202457]
152: 3434750 United_States
    [pem_pos=0, final_scores=0.47376937,
final_local_scores=0.44678026, pem=-0.14075686,
similarity=18.187986]
153: 1157765 Jon_Drummond
    [pem_pos=0, final_scores=0.48765022,
final_local_scores=0.6027642, pem=0.0, similarity=22.077398]
154: 3434750 United_States
```

[pem_pos=0, final_scores=0.47801477,
final_local_scores=0.45141774, pem=-0.14075686,
similarity=18.326376]
155: 13974630 Patrick_Stevens
[pem_pos=0, final_scores=0.6097918,
final_local_scores=0.7415706, pem=0.0, similarity=26.219658]
156: 3343 Belgium
[pem_pos=0, final_scores=0.71419185,
final_local_scores=0.7150365, pem=-0.24180195,
similarity=26.742786]
157: 681338 Michael_Marsh_(athlete)
[pem_pos=0, final_scores=0.6431633,
final_local_scores=0.7779676, pem=0.0, similarity=27.305817]
158: 3434750 United_States
[pem_pos=0, final_scores=0.47388387,
final_local_scores=0.44690531, pem=-0.14075686,
similarity=18.191715]
159: 2688112 Iván_García_(athlete)
[pem_pos=0, final_scores=0.5712668,
final_local_scores=0.72676986, pem=0.0, similarity=25.777975]
160: 5042481 Cuba
[pem_pos=0, final_scores=0.52978104,
final_local_scores=0.52400094, pem=-0.23907764,
similarity=21.027088]
161: 3343 Belgium
[pem_pos=0, final_scores=0.7233586,
final_local_scores=0.7250497, pem=-0.24180195,
similarity=27.0416]
162: 1676088 LaMont_Smith
[pem_pos=0, final_scores=0.5788206,
final_local_scores=0.74458253, pem=0.0, similarity=26.309542]
163: 3434750 United_States
[pem_pos=0, final_scores=0.48690456,
final_local_scores=0.4611284, pem=-0.14075686,
similarity=18.616163]
164: 613736 Svetlana_Masterkova
[pem_pos=0, final_scores=0.74195296,
final_local_scores=0.8696478, pem=0.0, similarity=30.041737]
165: 25391 Russia
[pem_pos=0, final_scores=0.79612654,
final_local_scores=0.8192543, pem=-0.20299491,
similarity=29.641811]
166: 66748 Maria_Mutola
[pem_pos=0, final_scores=0.6441229,
final_local_scores=0.720318, pem=0.0, similarity=25.58544]
167: 19301 Mozambique
[pem_pos=0, final_scores=0.6065213,
final_local_scores=0.6323632, pem=-0.36112452,
similarity=24.924543]
168: 21285670 Małgorzata_Rydz
[pem_pos=0, final_scores=0.62248737,
final_local_scores=0.76535404, pem=0.0, similarity=26.929403]

169: 22936 Poland
[pem_pos=0, final_scores=0.5868542,
final_local_scores=0.5831816, pem=-0.15333635,
similarity=22.326881]
170: 3343 Belgium
[pem_pos=0, final_scores=0.69263405,
final_local_scores=0.6914878, pem=-0.24180195,
similarity=26.040045]
171: 3343 Belgium
[pem_pos=0, final_scores=0.6766829,
final_local_scores=0.6740637, pem=-0.24180195,
similarity=25.5200751]
172: 14532 Italy
[pem_pos=0, final_scores=0.7651718,
final_local_scores=0.7920611, pem=-0.20612453,
similarity=28.847334]
173: 3343 Belgium
[pem_pos=0, final_scores=0.7369532,
final_local_scores=0.7398998, pem=-0.24180195,
similarity=27.484756]
174: 26667 Spain
[pem_pos=0, final_scores=0.6538587,
final_local_scores=0.65300316, pem=-0.18857051,
similarity=24.602106]
175: 2722641 Mary_Onyali-Omagbemi
[pem_pos=0, final_scores=0.7094194,
final_local_scores=0.8443988, pem=0.0, similarity=29.288258]
176: 21383 Nigeria
[pem_pos=0, final_scores=0.6507475,
final_local_scores=0.67134064, pem=-0.31455088,
similarity=25.834435]
177: 1031484 Inger_Miller
[pem_pos=0, final_scores=0.6018694,
final_local_scores=0.7209283, pem=0.0, similarity=25.603651]
178: 3434750 United_States
[pem_pos=0, final_scores=0.4901325,
final_local_scores=0.4646545, pem=-0.14075686,
similarity=18.721388]
179: 49242 Irina_Privalova
[pem_pos=0, final_scores=0.7272405,
final_local_scores=0.8759039, pem=0.0, similarity=30.228434]
180: 25391 Russia
[pem_pos=0, final_scores=0.81356436,
final_local_scores=0.83830255, pem=-0.20299491,
similarity=30.210249]
181: 25391 Russia
[pem_pos=0, final_scores=0.8379026,
final_local_scores=0.86488837, pem=-0.20299491,
similarity=31.003624]
182: 6812531 Marina_Trandenkova
[pem_pos=0, final_scores=0.43700016,
final_local_scores=0.5614345, pem=0.0, similarity=20.844038]

183: 25391 Russia
[pem_pos=0, final_scores=0.8455736,
final_local_scores=0.8732677, pem=-0.20299491,
similarity=31.253681]
184: 2511007 Chandra_Sturrup
[pem_pos=0, final_scores=0.81716484,
final_local_scores=0.9684358, pem=0.0, similarity=32.98977]
185: 3451 The_Bahamas
[pem_pos=0, final_scores=0.52948064,
final_local_scores=0.54275346, pem=-0.35113433,
similarity=22.196081]
186: 3434750 United_States
[pem_pos=0, final_scores=0.501633,
final_local_scores=0.47721702, pem=-0.14075686,
similarity=19.09628]
187: 3524541 Galina_Malchugina
[pem_pos=0, final_scores=0.70742697,
final_local_scores=0.859131, pem=0.0, similarity=29.727896]
188: 25391 Russia
[pem_pos=0, final_scores=0.7744468,
final_local_scores=0.79557246, pem=-0.20299491,
similarity=28.935099]
189: 66002 Cathy_Freeman
[pem_pos=0, final_scores=0.712742,
final_local_scores=0.80120593, pem=0.0, similarity=27.999296]
190: 4689264 Australia
[pem_pos=0, final_scores=0.7730287,
final_local_scores=0.78855777, pem=-0.17168477,
similarity=28.555498]
191: 66161 Marie-José_Pérec
[pem_pos=0, final_scores=0.7683379,
final_local_scores=0.88507277, pem=0.0, similarity=30.50205]
192: 5843419 France
[pem_pos=0, final_scores=0.8001438,
final_local_scores=0.81288207, pem=-0.16240834,
similarity=29.230936]
193: 2489403 Falilat_Ogunkoya
[pem_pos=0, final_scores=0.8055256,
final_local_scores=0.9517172, pem=0.0, similarity=32.490852]
194: 21383 Nigeria
[pem_pos=0, final_scores=0.64636344,
final_local_scores=0.66655177, pem=-0.31455088,
similarity=25.691525]
195: 3670385 Pauline_Davis-Thompson
[pem_pos=1, final_scores=0.4743858,
final_local_scores=0.6065295, pem=-0.0845574,
similarity=22.6496]
196: 3451 The_Bahamas
[pem_pos=0, final_scores=0.5501167,
final_local_scores=0.5652953, pem=-0.35113433,
similarity=22.868774]
197: 21383 Nigeria

```
[pem_pos=0, final_scores=0.64208347,
final_local_scores=0.66187656, pem=-0.31455088,
similarity=25.552006]
198: 3434750 United_States
    [pem_pos=0, final_scores=0.5198412,
final_local_scores=0.49710673, pem=-0.14075686,
similarity=19.689827]
199: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.6204477,
final_local_scores=0.62524444, pem=-0.27794555,
similarity=24.259766]
200: 3343 Belgium
    [pem_pos=0, final_scores=0.7211086,
final_local_scores=0.722592, pem=-0.24180195,
similarity=26.968256]
201: 673604 Daniel_Komen
    [pem_pos=0, final_scores=0.42863578,
final_local_scores=0.5020155, pem=0.0, similarity=19.070856]
202: 188171 Kenya
    [pem_pos=0, final_scores=0.5974317,
final_local_scores=0.6134772, pem=-0.27389407,
similarity=23.886574]
203: 3517130 Khalid_Boulami
    [pem_pos=0, final_scores=0.51143914,
final_local_scores=0.6483526, pem=0.0, similarity=23.437849]
204: 19291 Morocco
    [pem_pos=0, final_scores=0.64975876,
final_local_scores=0.66664165, pem=-0.29598162,
similarity=25.593224]
205: 3058737 Bob_Kennedy_(athlete)
    [pem_pos=1, final_scores=0.4385236,
final_local_scores=0.5547843, pem=-0.6074065,
similarity=23.948753]
206: 3434750 United_States
    [pem_pos=0, final_scores=0.50143164,
final_local_scores=0.47699708, pem=-0.14075686,
similarity=19.089714]
207: 19291 Morocco
    [pem_pos=0, final_scores=0.63330656,
final_local_scores=0.6486701, pem=-0.29598162,
similarity=25.056915]
208: 2736339 Tom_Nyariki
    [pem_pos=0, final_scores=0.35449272,
final_local_scores=0.51111513, pem=0.0, similarity=19.342407]
209: 188171 Kenya
    [pem_pos=0, final_scores=0.636879,
final_local_scores=0.6565673, pem=-0.27389407,
similarity=25.172468]
210: 492499 Noureddine_Morceli
    [pem_pos=0, final_scores=0.7569397,
final_local_scores=0.896187, pem=0.0, similarity=30.833721]
211: 358 Algeria
```

```
[pem_pos=0, final_scores=0.62043625,
final_local_scores=0.6440603, pem=-0.33694586,
similarity=25.142122]
212: 187749 Ethiopia
    [pem_pos=0, final_scores=0.5096639,
final_local_scores=0.5122667, pem=-0.23532115,
similarity=20.656487]
213: 188171 Kenya
    [pem_pos=0, final_scores=0.6491769,
final_local_scores=0.6700009, pem=-0.27389407,
similarity=25.573355]
214: 3908095 Lars_Riedel
    [pem_pos=0, final_scores=0.45092618,
final_local_scores=0.6093126, pem=0.0, similarity=22.272816]
215: 11867 Germany
    [pem_pos=0, final_scores=0.8646962,
final_local_scores=0.8851165, pem=-0.15529114,
similarity=31.347855]
216: 6781457 Anthony_Washington
    [pem_pos=0, final_scores=0.4681732,
final_local_scores=0.5993902, pem=0.0, similarity=21.976713]
217: 3434750 United_States
    [pem_pos=0, final_scores=0.5145405,
final_local_scores=0.4913165, pem=-0.14075686,
similarity=19.517036]
218: 3310705 Vladimir_Dubrovshchik
    [pem_pos=0, final_scores=0.767531,
final_local_scores=0.92138225, pem=0.0, similarity=31.585598]
219: 3457 Belarus
    [pem_pos=0, final_scores=0.64532024,
final_local_scores=0.64604986, pem=-0.4220667,
similarity=25.664392]
220: 17675 Lithuania
    [pem_pos=0, final_scores=0.6525554,
final_local_scores=0.65583855, pem=-0.39766282,
similarity=25.823795]
221: 3918291 Jürgen_Schult
    [pem_pos=0, final_scores=0.57299465,
final_local_scores=0.7146159, pem=0.0, similarity=25.415276]
222: 11867 Germany
    [pem_pos=0, final_scores=0.78998893,
final_local_scores=0.80351007, pem=-0.15529114,
similarity=28.912554]
223: 3457 Belarus
    [pem_pos=0, final_scores=0.6222295,
final_local_scores=0.62082666, pem=-0.4220667,
similarity=24.911682]
224: 17675 Lithuania
    [pem_pos=0, final_scores=0.6279741,
final_local_scores=0.62898713, pem=-0.39766282,
similarity=25.022495]
225: 3026211 Michael_Möllenbeck
```

[pem_pos=0, final_scores=0.47294295,
final_local_scores=0.648044, pem=0.0, similarity=23.428638]
226: 11867 Germany
[pem_pos=0, final_scores=0.8607331,
final_local_scores=0.8807873, pem=-0.15529114,
similarity=31.218662]
227: 52056 Jonathan_Edwards_(athlete)
[pem_pos=1, final_scores=0.41301113,
final_local_scores=0.47954923, pem=-1.4156067,
similarity=26.098707]
228: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5875271,
final_local_scores=0.57744163, pem=-0.61114657,
similarity=24.645233]
229: 3241060 Yoelbi_Quesada
[pem_pos=0, final_scores=0.60128224,
final_local_scores=0.76948994, pem=0.0, similarity=27.052828]
230: 5042481 Cuba
[pem_pos=0, final_scores=0.5388766,
final_local_scores=0.53393644, pem=-0.23907764,
similarity=21.323582]
231: 2690829 Brian_Wellman
[pem_pos=0, final_scores=0.54667926,
final_local_scores=0.7067333, pem=0.0, similarity=25.180044]
232: 3460 Bermuda
[pem_pos=0, final_scores=0.42321616,
final_local_scores=0.44760895, pem=-0.4078151,
similarity=19.665018]
233: 2645623 Kenny_Harrison
[pem_pos=0, final_scores=0.6771988,
final_local_scores=0.80747885, pem=0.0, similarity=28.186493]
234: 3434750 United_States
[pem_pos=0, final_scores=0.5087132,
final_local_scores=0.48495108, pem=-0.14075686,
similarity=19.32708]
235: 25391 Russia
[pem_pos=0, final_scores=0.83882517,
final_local_scores=0.86589605, pem=-0.20299491,
similarity=31.033695]
236: 5912006 Francis_Agyepong
[pem_pos=0, final_scores=0.63316834,
final_local_scores=0.8017518, pem=0.0, similarity=28.015585]
237: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5917193,
final_local_scores=0.58202106, pem=-0.61114657,
similarity=24.781893]
238: 9788550 Rogel_Nachum
[pem_pos=0, final_scores=0.6049545,
final_local_scores=0.78388906, pem=0.0, similarity=27.482527]
239: 9282173 Israel
[pem_pos=0, final_scores=0.4781255,
final_local_scores=0.4918552, pem=-0.21279564,

similarity=19.92487]
240: 8485751 Sigurd_Njerve
[pem_pos=0, final_scores=0.49562532,
final_local_scores=0.62757343, pem=0.0, similarity=22.817757]
241: 21241 Norway
[pem_pos=0, final_scores=0.7830474,
final_local_scores=0.8250643, pem=-0.22668563,
similarity=29.94403]
242: 492488 Hicham_El_Guerrouj
[pem_pos=0, final_scores=0.6173414,
final_local_scores=0.68540525, pem=0.0, similarity=24.543573]
243: 19291 Morocco
[pem_pos=0, final_scores=0.64047974,
final_local_scores=0.6565057, pem=-0.29598162,
similarity=25.290747]
244: 8903845 Isaac_Viciosa
[pem_pos=0, final_scores=0.5801851,
final_local_scores=0.7115275, pem=0.0, similarity=25.323114]
245: 26667 Spain
[pem_pos=0, final_scores=0.63139015,
final_local_scores=0.62845975, pem=-0.18857051,
similarity=23.86968]
246: 815988 William_Tanui
[pem_pos=0, final_scores=0.5006073,
final_local_scores=0.6325039, pem=0.0, similarity=22.964891]
247: 188171 Kenya
[pem_pos=0, final_scores=0.6451636,
final_local_scores=0.665617, pem=-0.27389407,
similarity=25.44253]
248: 188171 Kenya
[pem_pos=0, final_scores=0.66711825,
final_local_scores=0.68959904, pem=-0.27389407,
similarity=26.158201]
249: 14560 Republic_of_Ireland
[pem_pos=1, final_scores=0.5658335,
final_local_scores=0.5983022, pem=-1.5040284,
similarity=30.123386]
250: 13182495 John_Mayock
[pem_pos=0, final_scores=0.49818778,
final_local_scores=0.6382393, pem=0.0, similarity=23.136047]
251: 31717 United_Kingdom
[pem_pos=0, final_scores=0.577606,
final_local_scores=0.5666043, pem=-0.61114657,
similarity=24.321825]
252: 6261028 Laban_Rotich
[pem_pos=0, final_scores=0.6960351,
final_local_scores=0.84446996, pem=0.0, similarity=29.29038]
253: 188171 Kenya
[pem_pos=0, final_scores=0.59629875,
final_local_scores=0.6122396, pem=-0.27389407,
similarity=23.849642]
254: 3343 Belgium

[pem_pos=0, final_scores=0.7185269,
final_local_scores=0.7197718, pem=-0.24180195,
similarity=26.884096]
255: 826515 Stefka_Kostadinova
[pem_pos=0, final_scores=0.72425854,
final_local_scores=0.8402442, pem=0.0, similarity=29.164274]
256: 3415 Bulgaria
[pem_pos=0, final_scores=0.750869,
final_local_scores=0.7715109, pem=-0.31727514,
similarity=28.83853]
257: 6299154 Inha_Babakova
[pem_pos=0, final_scores=0.47488022,
final_local_scores=0.6644582, pem=0.0, similarity=23.918472]
258: 31750 Ukraine
[pem_pos=0, final_scores=0.67993313,
final_local_scores=0.68248713, pem=-0.35407752,
similarity=26.382019]
259: 3905854 Alina_Astafei
[pem_pos=0, final_scores=0.680673,
final_local_scores=0.84144247, pem=0.0, similarity=29.200035]
260: 11867 Germany
[pem_pos=0, final_scores=0.89278024,
final_local_scores=0.9157941, pem=-0.15529114,
similarity=32.263332]
261: 8278156 Tatyana_Babashkina
[pem_pos=0, final_scores=0.625369,
final_local_scores=0.8210625, pem=0.0, similarity=28.591856]
262: 25391 Russia
[pem_pos=0, final_scores=0.8274606,
final_local_scores=0.85348207, pem=-0.20299491,
similarity=30.663237]
263: 8277997 Svetlana_Zalevskaya
[pem_pos=0, final_scores=0.5561861,
final_local_scores=0.7197964, pem=0.0, similarity=25.569874]
264: 16642 Kazakhstan
[pem_pos=0, final_scores=0.734636,
final_local_scores=0.7515257, pem=-0.41320926,
similarity=28.763836]
265: 8278273 Yelena_Gulyayeva
[pem_pos=0, final_scores=0.64252394,
final_local_scores=0.8084968, pem=0.0, similarity=28.21687]
266: 25391 Russia
[pem_pos=0, final_scores=0.8239897,
final_local_scores=0.8496905, pem=-0.20299491,
similarity=30.55009]
267: 21241 Norway
[pem_pos=0, final_scores=0.7876921,
final_local_scores=0.83013797, pem=-0.22668563,
similarity=30.095438]
268: 19260 Moldova
[pem_pos=0, final_scores=0.5198591,
final_local_scores=0.5194753, pem=-0.47343826,

similarity=22.166523]
269: 10906630 Nelé Žilinskienė
[pem_pos=0, final_scores=0.48300213,
final_local_scores=0.6727516, pem=0.0, similarity=24.165962]
270: 17675 Lithuania
[pem_pos=0, final_scores=0.62136835,
final_local_scores=0.62177134, pem=-0.39766282,
similarity=24.807161]
271: 3580285 Salah_Hissou
[pem_pos=0, final_scores=0.5832602,
final_local_scores=0.73800474, pem=0.0, similarity=26.113247]
272: 19291 Morocco
[pem_pos=0, final_scores=0.64594024,
final_local_scores=0.6624704, pem=-0.29598162,
similarity=25.468744]
273: 597066 Paul_Tergat
[pem_pos=0, final_scores=0.7502963,
final_local_scores=0.83217853, pem=0.0, similarity=28.92358]
274: 188171 Kenya
[pem_pos=0, final_scores=0.645307,
final_local_scores=0.66577363, pem=-0.27389407,
similarity=25.447205]
275: 7080613 Paul_Koech
[pem_pos=0, final_scores=0.20800208,
final_local_scores=0.31939825, pem=0.0, similarity=13.621194]
276: 188171 Kenya
[pem_pos=0, final_scores=0.67387706,
final_local_scores=0.6969821, pem=-0.27389407,
similarity=26.378527]
277: 188171 Kenya
[pem_pos=0, final_scores=0.5932967,
final_local_scores=0.6089602, pem=-0.27389407,
similarity=23.751778]
278: 7190467 Aloës_Nizigama
[pem_pos=0, final_scores=0.49467927,
final_local_scores=0.65647143, pem=0.0, similarity=23.68013]
279: 21490998 Burundi
[pem_pos=0, final_scores=0.45231777,
final_local_scores=0.47811288, pem=-0.3738042,
similarity=20.39036]
280: 7190531 Mathias_Ntawulikura
[pem_pos=0, final_scores=0.6043072,
final_local_scores=0.7428859, pem=0.0, similarity=26.25891]
281: 25645 Rwanda
[pem_pos=0, final_scores=0.63948977,
final_local_scores=0.6612245, pem=-0.42949,
similarity=26.157604]
282: 3552171 Abel_Antón
[pem_pos=0, final_scores=0.69261354,
final_local_scores=0.84620386, pem=0.0, similarity=29.342125]
283: 26667 Spain
[pem_pos=0, final_scores=0.61699814,

```
final_local_scores=0.61273867, pem=-0.18857051,  
similarity=23.400532]  
284: 8380236 Kamiel_Maase  
    [pem_pos=0, final_scores=0.58433807,  
final_local_scores=0.74975055, pem=0.0, similarity=26.463764]  
285: 21148 Netherlands  
    [pem_pos=0, final_scores=0.6266925,  
final_local_scores=0.6294647, pem=-0.24940795,  
similarity=24.230515]  
286: 187749 Ethiopia  
    [pem_pos=0, final_scores=0.54567367,  
final_local_scores=0.55160195, pem=-0.23532115,  
similarity=21.830328]  
287: 26830 Slovakia  
    [pem_pos=0, final_scores=0.6602095,  
final_local_scores=0.6699299, pem=-0.36602598,  
similarity=26.072266]
```

Sample: b'162_SOCCER'

SOCCKER - JORGE CALLS UP SIX [1 1 PORTO 1] PLAYERS FOR [1 1 WO RLD CUP 1] QUALIFIER . [2 2 LISBON 2] 1996-08-23 [3 3 Portugal 3] 's new coach [4 4 Artur Jorge 4] called up six players from league champions [5 5 Porto 5] on Friday in an 18-man squad for the opening [6 6 World Cup 6] qualifier against [7 7 Armenia 7] on August 31 . Midfielder [8 8 Paulo Sousa 8] , recently transferred to [9 9 Borussia Dortmund 9] from [10 10 Italy 10] 's [11 11 Juventus 11] , is the only leading member of the [12 12 Portuguese 12] side from this year 's [13 13 European 13] championships who will not make the trip . It will be Jorge 's first game in charge of the national squad since taking over from [14 14 Antonio Oliveira 14] , who now coaches [15 15 Porto 15] , at the end of [16 16 Euro 96 16] . Squad : Goalkeepers - [17 17 Vitor Baia 17] , [18 18 Rui Correia 18] . Defenders - [19 19 Jorge Costa 19] , [20 20 Paulinho Santos 20] , [21 21 Helder Cristovao 21] , [22 22 Carlos Secreto 22] , [23 23 Dimas Teixeira 23] , [24 24 Fernando Couto 24] . Midfielders - Jose Barroso , [25 25 Luis Figo 25] , [26 26 Rui Barros 26] , [27 27 Rui Costa 27] , [28 28 Oceano Cruz 28] , [29 29 Ricardo Sa Pinto 29] . Forwards - Domingos Oliveira , [30 30 Joao Vieira Pinto 30] , [31 31 Jorge Cadete 31] , [32 32 Antonio Folha 32] .

1: 172589 F.C._Porto
2: 18091 Lisbon
3: 576328 Portugal_national_football_team
4: 2695399 Artur_Jorge_(footballer)
5: 172589 F.C._Porto
6: 126460 1998_FIFA_World_Cup
7: 695292 Armenia_national_football_team
8: 1469201 Paulo_Sousa
9: 331715 Borussia_Dortmund
10: 14532 Italy
11: 16415 Juventus_F.C.

```
12: 23033 Portugal
13: 9239 Europe
14: 1965301 António_Luís_Alves_Ribeiro_Oliveira
15: 172589 F.C._Porto
16: 555876 UEFA_Euro_1996
17: 670993 Vítor_Baía
18: 9390856 Rui_Correia
19: 670891 Jorge_Costa
20: 7595668 Paulinho_Santos
21: 4935997 Hélder_Cristóvão
22: 5571094 Carlos_Secretário
23: 2135282 Dimas_Teixeira
24: 772771 Fernando_Couto
25: 353781 Luís_Figo
26: 3381791 Rui_Barros
27: 612521 Rui_Costa
28: 3304564 Oceano_da_Cruz
29: 3933665 Ricardo_Sá_Pinto
30: 1254376 João_Vieira_Pinto
31: 3526090 Jorge_Cadete
32: 7611575 António_Folha
1: 126460 1998_FIFA_World_Cup
    [winner: 11370 FIFA_World_Cup, pem_pos=0,
final_scores=0.102128014, final_local_scores=0.058450118,
pem=-1.001087, similarity=11.278048]
    gt: 126460 1998_FIFA_World_Cup, pem_pos=18,
final_scores=0.09241508, final_local_scores=0.034399748,
pem=-4.2172847, similarity=28.05052 ]
1: 172589 F.C._Porto
    [pem_pos=1, final_scores=0.7424651,
final_local_scores=0.73089254, pem=-1.1013846,
similarity=31.890505]
2: 18091 Lisbon
    [pem_pos=0, final_scores=0.4740134,
final_local_scores=0.5011605, pem=-0.09749677,
similarity=19.575544]
3: 576328 Portugal_national_football_team
    [pem_pos=1, final_scores=0.45210737,
final_local_scores=0.39669907, pem=-2.453637,
similarity=29.271263]
4: 2695399 Artur_Jorge_(footballer)
    [pem_pos=0, final_scores=0.5404857,
final_local_scores=0.5749436, pem=0.0, similarity=21.247177]
5: 172589 F.C._Porto
    [pem_pos=1, final_scores=0.63655716,
final_local_scores=0.6152042, pem=-1.1013846,
similarity=28.438135]
6: 126460 1998_FIFA_World_Cup
    [pem_pos=18, final_scores=0.36375153,
final_local_scores=0.32742214, pem=-4.2172847,
similarity=36.79489]
7: 695292 Armenia_national_football_team
```

```
[pem_pos=1, final_scores=0.42470378,
final_local_scores=0.44244623, pem=-2.6870568,
similarity=31.905825]
8: 1469201 Paulo_Sousa
    [pem_pos=0, final_scores=0.35507333,
final_local_scores=0.35319248, pem=0.0, similarity=14.629681]
9: 331715 Borussia_Dortmund
    [pem_pos=0, final_scores=0.46691656,
final_local_scores=0.4589749, pem=0.0, similarity=17.78644]
10: 14532 Italy
    [pem_pos=0, final_scores=0.6330084,
final_local_scores=0.7260805, pem=-0.20612453,
similarity=26.87834]
11: 16415 Juventus_F.C.
    [pem_pos=0, final_scores=0.4429162,
final_local_scores=0.40877718, pem=0.0, similarity=16.288439]
12: 23033 Portugal
    [pem_pos=1, final_scores=0.53750706,
final_local_scores=0.5729513, pem=-1.127839,
similarity=27.321087]
13: 9239 Europe
    [pem_pos=0, final_scores=0.5264275,
final_local_scores=0.63743615, pem=-0.25557294,
similarity=24.501925]
14: 1965301 António_Luís_Alves_Ribeiro_Oliveira
    [pem_pos=0, final_scores=0.39111757,
final_local_scores=0.37925953, pem=-0.6931472,
similarity=19.177015]
15: 172589 F.C._Porto
    [pem_pos=1, final_scores=0.62616485,
final_local_scores=0.6038522, pem=-1.1013846,
similarity=28.09937]
16: 555876 UEFA_Euro_1996
    [pem_pos=0, final_scores=0.48180705,
final_local_scores=0.4607504, pem=0.0, similarity=17.839424]
17: 670993 Vítor_Baía
    [pem_pos=0, final_scores=0.7497623,
final_local_scores=0.7647006, pem=0.0, similarity=26.909904]
18: 9390856 Rui_Correia
    [pem_pos=0, final_scores=0.47456992,
final_local_scores=0.5415164, pem=0.0, similarity=20.249643]
19: 670891 Jorge_Costa
    [pem_pos=0, final_scores=0.5092049,
final_local_scores=0.50389457, pem=0.0, similarity=19.126932]
20: 7595668 Paulinho_Santos
    [pem_pos=0, final_scores=0.4903878,
final_local_scores=0.51287544, pem=0.0, similarity=19.39494]
21: 4935997 Hélder_Cristóvão
    [pem_pos=0, final_scores=0.6808613,
final_local_scores=0.7448864, pem=0.0, similarity=26.31861]
22: 5571094 Carlos_Secretário
    [pem_pos=0, final_scores=0.531955,
```

```
final_local_scores=0.55856043, pem=0.0, similarity=20.75827]
23: 2135282 Dimas_Teixeira
    [pem_pos=0, final_scores=0.59951806,
final_local_scores=0.65254015, pem=0.0, similarity=23.562813]
24: 772771 Fernando_Couto
    [pem_pos=0, final_scores=0.74882764,
final_local_scores=0.76341295, pem=0.0, similarity=26.871479]
25: 353781 Luís_Figo
    [pem_pos=0, final_scores=0.6702817,
final_local_scores=0.68363345, pem=0.0, similarity=24.490698]
26: 3381791 Rui_Barros
    [pem_pos=0, final_scores=0.59618837,
final_local_scores=0.6203144, pem=0.0, similarity=22.601133]
27: 612521 Rui_Costa
    [pem_pos=0, final_scores=0.6797386,
final_local_scores=0.6957726, pem=0.0, similarity=24.852955]
28: 3304564 Oceano_da_Cruz
    [pem_pos=0, final_scores=0.5878702,
final_local_scores=0.6084027, pem=0.0, similarity=22.245665]
29: 3933665 Ricardo_Sá_Pinto
    [pem_pos=0, final_scores=0.5943721,
final_local_scores=0.61022615, pem=0.0, similarity=22.30008]
30: 1254376 João_Vieira_Pinto
    [pem_pos=0, final_scores=0.32699198,
final_local_scores=0.31504315, pem=0.0, similarity=13.491229]
31: 3526090 Jorge_Cadete
    [pem_pos=0, final_scores=0.42917973,
final_local_scores=0.44567168, pem=0.0, similarity=17.389444]
32: 7611575 António_Folha
    [pem_pos=0, final_scores=0.45134747,
final_local_scores=0.49738735, pem=0.0, similarity=18.932743]
1: 11370 FIFA_World_Cup
    [pem_pos=0, final_scores=0.102128014,
final_local_scores=0.058450118, pem=-1.001087,
similarity=11.278048]
```

Sample: b'163_SOCCER'

```
SOCCER - [1 1 VOGTS 1] KEEPS FAITH WITH EURO ' 96
CHAMPIONS . [2 2 BONN 2] 1996-08-23 Trainer [3 3 Berti Vogts 3
] kept faith with his entire [4 4 European 4] championship
winning squad for [5 5 Germany 5] 's first match since their
title victory , a friendly in [6 6 Poland 6] . [7 7 Vogts 7]
picked no new players for the squad for the September 4 game
in [8 8 Zabrze 8] . Instead on Friday he nominated all 23
Euro ' 96 veterans including [9 9 Bremen 9] 's [10 10 Jens To
dt 10] , called up before the final by special [11 11 UEFA 11]
dispensation . He will , however , have to do without the [12
12 Dortmund 12] trio of libero [13 13 Matthias Sammer 13] ,
midfielder [14 14 Steffen Freund 14] and defender [15 15 Rene
Schneider 15] , who were all formally nominated despite being
injured . " This squad is currently the basis of my planning
for the 1998 [16 16 World Cup 16] , " [17 17 Vogts 17] said .
```

" We 'll have to see which other players produce good league performances to play themselves into the squad . " Squad : Goalkeepers - [18 18 Oliver Kahn 18] , [19 19 Andreas Koepke 19] , [20 20 Oliver Reck 20] Defenders - [21 21 Markus Babbel 21] , [22 22 Thomas Helmer 22] , [23 23 Juergen Kohler 23] , [24 24 Stefan Reuter 24] , [25 25 Matthias Sammer 25] , [26 26 René Schneider 26] Midfielders - [27 27 Mario Basler 27] , [28 28 Marco Bode 28] , [29 29 Dieter Eilts 29] , [30 30 Steffen Freund 30] , [31 31 Thomas Haessler 31] , [32 32 Andreas Möller 32] , [33 33 Mehmet Scholl 33] , [34 34 Thomas Strunz 34] , [35 35 Jens Todt 35] , [36 36 Christian Ziege 36] Forwards - [37 37 Oliver Bierhoff 37] , [38 38 Fredi Bobic 38] , [39 39 Jürgen Klinsmann 39] , [40 40 Stefan Kuntz 40] .

- 1: 775222 Berti_Vogts
- 2: 3295 Bonn
- 3: 775222 Berti_Vogts
- 4: 9239 Europe
- 5: 250204 Germany_national_football_team
- 6: 22936 Poland
- 7: 775222 Berti_Vogts
- 8: 72873 Zabrze
- 9: 7598082 SV_Werder_Bremen
- 10: 10104303 Jens_Todt
- 11: 32332 UEFA
- 12: 331715 Borussia_Dortmund
- 13: 971550 Matthias_Sammer
- 14: 3640696 Steffen_Freund
- 15: 5659152 René_Schneider_(footballer)
- 16: 11370 FIFA_World_Cup
- 17: 775222 Berti_Vogts
- 18: 347547 Oliver_Kahn
- 19: 2255217 Andreas_Köpke
- 20: 5659106 Oliver_Reck
- 21: 1265241 Markus_Babbel
- 22: 2568262 Thomas_Helmer
- 23: 3428316 Jürgen_Kohler
- 24: 5104668 Stefan_Reuter
- 25: 971550 Matthias_Sammer
- 26: 5659152 René_Schneider_(footballer)
- 27: 3272105 Mario_Basler
- 28: 3524294 Marco_Bode
- 29: 3631581 Dieter_Eilts
- 30: 3640696 Steffen_Freund
- 31: 1462547 Thomas_Häßler
- 32: 1462717 Andreas_Möller
- 33: 1120394 Mehmet_Scholl
- 34: 1681297 Thomas_Strunz
- 35: 10104303 Jens_Todt
- 36: 1462512 Christian_Ziege
- 37: 1129681 Oliver_Bierhoff
- 38: 2194769 Fredi_Bobic
- 39: 525141 Jürgen_Klinsmann

40: 3660396 Stefan_Kuntz
1: 775222 Berti_Vogts
[pem_pos=0, final_scores=0.6236921,
final_local_scores=0.6186947, pem=0.0, similarity=22.552797]
2: 3295 Bonn
[pem_pos=0, final_scores=0.38566142,
final_local_scores=0.45560116, pem=0.0, similarity=17.68576]
3: 775222 Berti_Vogts
[pem_pos=0, final_scores=1.0089061,
final_local_scores=1.0394826, pem=0.0, similarity=35.109947]
4: 9239 Europe
[pem_pos=0, final_scores=0.4792568,
final_local_scores=0.6098444, pem=-0.25557294,
similarity=23.678532]
5: 250204 Germany_national_football_team
[pem_pos=1, final_scores=0.54294026,
final_local_scores=0.50765085, pem=-3.3613694,
similarity=37.518677]
6: 22936 Poland
[pem_pos=0, final_scores=0.26098686,
final_local_scores=0.32926026, pem=-0.15333635,
similarity=14.749363]
7: 775222 Berti_Vogts
[pem_pos=0, final_scores=0.6443004,
final_local_scores=0.6412062, pem=0.0, similarity=23.224585]
8: 72873 Zabrze
[pem_pos=0, final_scores=0.42117298,
final_local_scores=0.51632476, pem=0.0, similarity=19.497873]
9: 7598082 SV_Werder_Bremen
[pem_pos=10, final_scores=0.6804256,
final_local_scores=0.659603, pem=-3.3898606,
similarity=42.20817]
10: 10104303 Jens_Todt
[pem_pos=0, final_scores=0.5482127,
final_local_scores=0.602106, pem=0.0, similarity=22.057756]
11: 32332 UEFA
[pem_pos=0, final_scores=0.30362403,
final_local_scores=0.30614346, pem=0.0, similarity=13.225644]
12: 331715 Borussia_Dortmund
[pem_pos=2, final_scores=0.47262436,
final_local_scores=0.43160993, pem=-2.1086495,
similarity=28.43698]
13: 971550 Matthias_Sammer
[pem_pos=0, final_scores=0.882168,
final_local_scores=0.9059529, pem=0.0, similarity=31.125153]
14: 3640696 Steffen_Freund
[pem_pos=0, final_scores=0.80185527,
final_local_scores=0.85326546, pem=0.0, similarity=29.552856]
15: 5659152 René_Schneider_(footballer)
[pem_pos=1, final_scores=0.4571715,
final_local_scores=0.5443495, pem=-0.6931472,
similarity=24.10363]

```
16: 11370 FIFA_World_Cup
    [pem_pos=0, final_scores=0.20820032,
final_local_scores=0.19226933, pem=-1.001087,
similarity=15.27148]
17: 775222 Berti_Vogts
    [pem_pos=0, final_scores=0.67927784,
final_local_scores=0.6794137, pem=0.0, similarity=24.364773]
18: 347547 Oliver_Kahn
    [pem_pos=0, final_scores=0.7192921,
final_local_scores=0.7353925, pem=0.0, similarity=26.035292]
19: 2255217 Andreas_Köpke
    [pem_pos=0, final_scores=1.0528784,
final_local_scores=1.1195693, pem=0.0, similarity=37.499893]
20: 5659106 Oliver_Reck
    [pem_pos=0, final_scores=0.6459364,
final_local_scores=0.705901, pem=0.0, similarity=25.155207]
21: 1265241 Markus_Babbel
    [pem_pos=0, final_scores=0.79936546,
final_local_scores=0.83793366, pem=0.0, similarity=29.095325]
22: 2568262 Thomas_Helmer
    [pem_pos=0, final_scores=0.8866325,
final_local_scores=0.93756837, pem=0.0, similarity=32.068623]
23: 3428316 Jürgen_Kohler
    [pem_pos=0, final_scores=0.69907135,
final_local_scores=0.75391376, pem=0.0, similarity=26.588003]
24: 5104668 Stefan_Reuter
    [pem_pos=0, final_scores=0.8236831,
final_local_scores=0.88358146, pem=0.0, similarity=30.457546]
25: 971550 Matthias_Sammer
    [pem_pos=0, final_scores=0.93314916,
final_local_scores=0.9616421, pem=0.0, similarity=32.78703]
26: 5659152 René_Schneider_(footballer)
    [pem_pos=1, final_scores=0.57995677,
final_local_scores=0.6784738, pem=-0.6931472,
similarity=28.106165]
27: 3272105 Mario_Basler
    [pem_pos=0, final_scores=0.6721738,
final_local_scores=0.700385, pem=0.0, similarity=24.990597]
28: 3524294 Marco_Bode
    [pem_pos=0, final_scores=0.7915731,
final_local_scores=0.8548199, pem=0.0, similarity=29.599243]
29: 3631581 Dieter_Eilts
    [pem_pos=0, final_scores=0.9502607,
final_local_scores=1.0347419, pem=0.0, similarity=34.968475]
30: 3640696 Steffen_Freund
    [pem_pos=0, final_scores=0.75114566,
final_local_scores=0.7978729, pem=0.0, similarity=27.899832]
31: 1462547 Thomas_Häßler
    [pem_pos=0, final_scores=0.88893217,
final_local_scores=0.9385969, pem=0.0, similarity=32.099316]
32: 1462717 Andreas_Möller
    [pem_pos=0, final_scores=0.821253,
```

```
final_local_scores=0.8863918, pem=0.0, similarity=30.541412]
33: 1120394 Mehmet_Scholl
    [pem_pos=0, final_scores=0.7855223,
final_local_scores=0.8191054, pem=0.0, similarity=28.533451]
34: 1681297 Thomas_Strunz
    [pem_pos=0, final_scores=0.80986124,
final_local_scores=0.86463267, pem=0.0, similarity=29.892075]
35: 10104303 Jens_Todt
    [pem_pos=0, final_scores=0.6146025,
final_local_scores=0.67462665, pem=0.0, similarity=24.221918]
36: 1462512 Christian_Ziege
    [pem_pos=0, final_scores=0.82013613,
final_local_scores=0.8566051, pem=0.0, similarity=29.652517]
37: 1129681 Oliver_Bierhoff
    [pem_pos=0, final_scores=0.8996882,
final_local_scores=0.94639, pem=0.0, similarity=32.33188]
38: 2194769 Fredi_Bobic
    [pem_pos=0, final_scores=0.73147655,
final_local_scores=0.7931702, pem=0.0, similarity=27.759495]
39: 525141 Jürgen_Klinsmann
    [pem_pos=0, final_scores=0.84489304,
final_local_scores=0.8478693, pem=0.0, similarity=29.391823]
40: 3660396 Stefan_Kuntz
    [pem_pos=0, final_scores=0.77022845,
final_local_scores=0.8426416, pem=0.0, similarity=29.235819]
```

Sample: b'164_SOCCER'

SOCCER - EUROPEAN CUP DRAWS FOR [1 1 AEK 1] , [2 2 OLYMPIAKOS 2] , 1 PAO . [3 3 ATHENS 3] 1996-08-23 Following are the [4 4 European 4] soccer draws for the [5 5 UEFA 5] cup and the cup 's winners cup involving [6 6 Greek 6] teams that took place today in [7 7 Geneva 7] : x-AEK Athens ([8 8 Greece 8]) v Chemlon Humenne ([9 9 Slovakia 9]) x-Olympiakos v [10 10 Ferencvaros 10] ([11 11 Hungary 11]) x-PAO v [12 12 Legia Warsaw 12] ([13 13 Poland 13]) x indicates seeded teams . -- Dimitris Kontogiannis , Athens Newsroom +301 3311812-4

```
1: 990655 A.E.K._Athens_F.C.
2: 1229550 Olympiacos_F.C.
3: 1216 Athens
4: 9239 Europe
5: 32332 UEFA
6: 12108 Greece
7: 12521 Geneva
8: 12108 Greece
9: 26830 Slovakia
10: 861261 Ferencvárosi_TC
11: 13275 Hungary
12: 677536 Legia_Warsaw
13: 22936 Poland
1: 1053145 Panathinaikos_F.C.
    [winner: 4772097 Nissan_Pao, pem_pos=0,
final_scores=-0.46645105, final_local_scores=-0.32378298,
```

```
pem=-0.0777826, similarity=-5.1496234|
    RECALL MISS: 1053145 Panathinaikos_F.C.]
1: 990655 A.E.K._Athens_F.C.
    [pem_pos=0, final_scores=0.25887257,
final_local_scores=0.25282973, pem=0.0, similarity=11.634656]
2: 1229550 Olympiacos_F.C.
    [pem_pos=0, final_scores=0.2676801,
final_local_scores=0.26787826, pem=-0.4098014,
similarity=14.312299]
3: 1216 Athens
    [pem_pos=0, final_scores=0.45630533,
final_local_scores=0.4546497, pem=-0.28831622,
similarity=19.225275]
4: 9239 Europe
    [pem_pos=0, final_scores=0.56722826,
final_local_scores=0.593973, pem=-0.25557294,
similarity=23.204897]
5: 32332 UEFA
    [pem_pos=0, final_scores=0.29299337,
final_local_scores=0.29015568, pem=0.0, similarity=12.748537]
6: 12108 Greece
    [pem_pos=1, final_scores=0.80162865,
final_local_scores=0.8038075, pem=-2.3247876,
similarity=40.71949]
7: 12521 Geneva
    [pem_pos=0, final_scores=0.48086977,
final_local_scores=0.52273047, pem=-0.24772356,
similarity=21.036192]
8: 12108 Greece
    [pem_pos=0, final_scores=0.8951797,
final_local_scores=0.9059978, pem=-0.27607638,
similarity=32.627842]
9: 26830 Slovakia
    [pem_pos=0, final_scores=0.52971524,
final_local_scores=0.5095859, pem=-0.36602598,
similarity=21.28728]
10: 861261 Ferencvárosi_TC
    [pem_pos=0, final_scores=0.30860645,
final_local_scores=0.31684464, pem=0.0, similarity=13.544989]
11: 13275 Hungary
    [pem_pos=0, final_scores=0.54450274,
final_local_scores=0.5230931, pem=-0.32577527,
similarity=21.471472]
12: 677536 Legia_Warsaw
    [pem_pos=0, final_scores=0.35591483,
final_local_scores=0.38659173, pem=0.0, similarity=15.626381]
13: 22936 Poland
    [pem_pos=0, final_scores=0.55252767,
final_local_scores=0.54383326, pem=-0.15333635,
similarity=21.152649]
```

Sample: b'165_SOCCER'

SOCCKER - [1 1 EURO 1] CLUB COMPETITION FIRST ROUND DRAWS . [2 2 GENEVA 2] 1996-08-23 Draws for the first round of the [3 3 E uropean 3] club soccer competitions made on Friday (x denotes seeded team) : [4 4 UEFA Cup 4] [5 5 Lyngby 5] ([6 6 Denmark 6]) v x-Club Brugge ([7 7 Belgium 7]) Casino Graz ([8 8 Austria 8]) v Ekeren ([9 9 Belgium 9]) [10 10 Besik tas 10] (Turkey) v [11 11 Molenbeek 11] ([12 12 Belgium 12]) [13 13 Alania Vladikavkaz 13] ([14 14 Russia 14]) v x- Anderlecht ([15 15 Belgium 15]) Cup Winners ' Cup x-Cercle Brugge ([16 16 Belgium 16]) v [17 17 Brann Bergen 17] ([18 18 Norway 18])

1: 9239 Europe
2: 12521 Geneva
3: 9239 Europe
4: 232175 UEFA_Europa_League
5: 1117976 Lyngby-Boldklub
6: 76972 Denmark
7: 3343 Belgium
8: 26964606 Austria
9: 3343 Belgium
10: 1216600 Beşiktaş_J.K.
11: 1452159 R.White_Daring_Molenbeek
12: 3343 Belgium
13: 1167720 FC_Alania_Vladikavkaz
14: 25391 Russia
15: 3343 Belgium
16: 3343 Belgium
17: 192068 SK_Brann
18: 21241 Norway
1: 9239 Europe
[pem_pos=1, final_scores=0.3720575,
final_local_scores=0.37224203, pem=-2.1768987,
similarity=27.03647]
2: 12521 Geneva
[pem_pos=0, final_scores=0.48289877,
final_local_scores=0.4981286, pem=-0.24772356,
similarity=20.302023]
3: 9239 Europe
[pem_pos=0, final_scores=0.51144594,
final_local_scores=0.52450275, pem=-0.25557294,
similarity=21.131765]
4: 232175 UEFA_Europa_League
[pem_pos=0, final_scores=0.31620592,
final_local_scores=0.3480868, pem=-0.260892,
similarity=15.89609]
5: 1117976 Lyngby-Boldklub
[pem_pos=1, final_scores=0.4201225,
final_local_scores=0.47597736, pem=-0.24484192,
similarity=19.625315]
6: 76972 Denmark
[pem_pos=0, final_scores=0.49068362,
final_local_scores=0.46483648, pem=-0.26762357,

```
similarity=19.416739]
7: 3343 Belgium
    [pem_pos=0, final_scores=0.5367483,
final_local_scores=0.4913491, pem=-0.24180195,
similarity=20.067507]
8: 26964606 Austria
    [pem_pos=0, final_scores=0.5091283,
final_local_scores=0.4827358, pem=-0.25262418,
similarity=19.869322]
9: 3343 Belgium
    [pem_pos=0, final_scores=0.5828221,
final_local_scores=0.5416777, pem=-0.24180195,
similarity=21.569414]
10: 1216600 Beşiktaş_J.K.
    [pem_pos=0, final_scores=0.24633227,
final_local_scores=0.2720245, pem=-0.09800805,
similarity=12.74045]
11: 1452159 R._White_Daring_Molenbeek
    [pem_pos=0, final_scores=0.37514472,
final_local_scores=0.42421365, pem=-0.005970144,
similarity=16.78156]
12: 3343 Belgium
    [pem_pos=0, final_scores=0.55360895,
final_local_scores=0.5097668, pem=-0.24180195,
similarity=20.617128]
13: 1167720 FC_Alania_Vladikavkaz
    [pem_pos=0, final_scores=0.23086394,
final_local_scores=0.29610667, pem=0.0, similarity=12.9261265]
14: 25391 Russia
    [pem_pos=0, final_scores=0.7107015,
final_local_scores=0.7498813, pem=-0.20299491,
similarity=27.571585]
15: 3343 Belgium
    [pem_pos=0, final_scores=0.4986301,
final_local_scores=0.4497108, pem=-0.24180195,
similarity=18.824938]
16: 3343 Belgium
    [pem_pos=0, final_scores=0.5708321,
final_local_scores=0.5285804, pem=-0.24180195,
similarity=21.178564]
17: 192068 SK_Brann
    [pem_pos=0, final_scores=0.54121834,
final_local_scores=0.60557777, pem=0.0, similarity=22.161362]
18: 21241 Norway
    [pem_pos=0, final_scores=0.57506555,
final_local_scores=0.58171606, pem=-0.22668563,
similarity=22.682032]
```

Sample: b'166_CRICKET'

CRICKET - [1 1 SRI LANKA 1] AND [2 2 AUSTRALIA 2] SAY
RELATIONS HAVE HEALED . [3 3 COLOMBO 3] 1996-08-23 [4 4 Sri La
nka 4] and [5 5 Australia 5] agreed on Friday that relations

between the two teams had healed since the [6 6 Sri Lankans 6] ' acrimonious tour last year . The [7 7 Sri Lankans 7] were first found guilty then cleared of ball tampering and off-spinner [8 8 Muttiah Muralitharan 8] was called for throwing during a controversial three-test series in [9 9 Australia 9] . " Our concern is to get out there and play proper cricket , " [10 10 Sri Lanka 10] captain [11 11 Arjuna Ranatunga 11] told a news conference on the eve of a warmup match between the [12 12 World Cup 12] champions and a [13 13 World XI 13] team scheduled for Saturday . " What happened is history . " [14 14 Australian 14] team manager Cam Battersby said he agreed with [15 15 Ranatunga 15] . " I believe relations between the two teams will be excellent , " Batterby said . The [16 16 Australians 16] are making their first visit to the Indian Ocean island since boycotting a [17 17 World Cup 17] fixture in February after a terrorist bomb in [18 18 Colombo 18] . [19 19 Australia 19] have been promised the presence of commandos , sniffer dogs and plainclothes policemen to ensure a limited overs tournament is trouble-free . The tournament , starting on August 26 , also includes [20 20 India 20] and [21 21 Zimbabwe 21] . Battersby said he was satisfied with the security arrangements . [22 22 Sri Lankan 22] officials said they expected heavy rain which washed out a warmup match on Friday should cease by Saturday . [23 23 Australia 23] , led by wicketkeeper [24 24 Ian Healy 24] , opened their short tour of [25 25 Sri Lanka 25] with a five-run win over the country 's youth team on Thursday .

1: 26750 Sri_Lanka
2: 4689264 Australia
3: 56636 Colombo
4: 397246 Sri_Lanka_national_cricket_team
5: 374833 Australia_national_cricket_team
6: 397246 Sri_Lanka_national_cricket_team
7: 397246 Sri_Lanka_national_cricket_team
8: 20905 Muttiah_Muralitharan
9: 4689264 Australia
10: 397246 Sri_Lanka_national_cricket_team
11: 1160705 Arjuna_Ranatunga
12: 7239 Cricket_World_Cup
13: 23538966 World_Series_Cricket_World_XI
14: 374833 Australia_national_cricket_team
15: 1160705 Arjuna_Ranatunga
16: 374833 Australia_national_cricket_team
17: 7239 Cricket_World_Cup
18: 56636 Colombo
19: 374833 Australia_national_cricket_team
20: 407754 India_national_cricket_team
21: 759317 Zimbabwe_national_cricket_team
22: 397246 Sri_Lanka_national_cricket_team
23: 374833 Australia_national_cricket_team
24: 928267 Ian_Healy
25: 26750 Sri_Lanka

```
1: 26750 Sri_Lanka
    [pem_pos=0, final_scores=0.469284,
final_local_scores=0.4944355, pem=-0.3027737,
similarity=20.491184]
2: 4689264 Australia
    [pem_pos=0, final_scores=0.6006995,
final_local_scores=0.67460674, pem=-0.17168477,
similarity=25.154972]
3: 56636 Colombo
    [pem_pos=0, final_scores=0.5603809,
final_local_scores=0.5794487, pem=-0.19393882,
similarity=22.436287]
4: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.59801763,
final_local_scores=0.5514074, pem=-2.2667584,
similarity=32.871796]
5: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.73213655,
final_local_scores=0.69425464, pem=-3.4989398,
similarity=43.835434]
6: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=0, final_scores=0.7816965,
final_local_scores=0.7520486, pem=-0.17988552,
similarity=27.51059]
7: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=0, final_scores=0.8204935,
final_local_scores=0.7944286, pem=-0.17988552,
similarity=28.775291]
8: 20905 Muttiah_Muralitharan
    [pem_pos=0, final_scores=0.69225603,
final_local_scores=0.70511466, pem=0.0, similarity=25.13174]
9: 4689264 Australia
    [pem_pos=0, final_scores=0.4348684,
final_local_scores=0.4934615, pem=-0.17168477,
similarity=19.749237]
10: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.66469496,
final_local_scores=0.62424225, pem=-2.2667584,
similarity=35.04533]
11: 1160705 Arjuna_Ranatunga
    [pem_pos=0, final_scores=0.61047006,
final_local_scores=0.65285623, pem=0.0, similarity=23.572245]
12: 7239 Cricket_World_Cup
    [pem_pos=4, final_scores=0.5495981,
final_local_scores=0.5367043, pem=-3.1731894,
similarity=37.362335]
13: 23538966 World_Series_Cricket_World_XI
    [pem_pos=0, final_scores=0.500681,
final_local_scores=0.5934069, pem=0.0, similarity=21.798159]
14: 374833 Australia_national_cricket_team
    [pem_pos=2, final_scores=0.7314995,
final_local_scores=0.6935588, pem=-3.6253178,
```

```
similarity=44.501934]
15: 1160705 Arjuna_Ranatunga
    [pem_pos=0, final_scores=0.36842638,
final_local_scores=0.38846028, pem=0.0, similarity=15.682143]
16: 374833 Australia_national_cricket_team
    [pem_pos=2, final_scores=0.5747267,
final_local_scores=0.5223083, pem=-1.4598138,
similarity=27.615128]
17: 7239 Cricket_World_Cup
    [pem_pos=4, final_scores=0.42480022,
final_local_scores=0.39798152, pem=-3.1731894,
similarity=33.222572]
18: 56636 Colombo
    [pem_pos=0, final_scores=0.5336563,
final_local_scores=0.55025613, pem=-0.19393882,
similarity=21.565123]
19: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.6788495,
final_local_scores=0.63604665, pem=-3.4989398,
similarity=42.098392]
20: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.6520689,
final_local_scores=0.61094975, pem=-3.5152702,
similarity=41.43826]
21: 759317 Zimbabwe_national_cricket_team
    [pem_pos=1, final_scores=0.7462319,
final_local_scores=0.7259672, pem=-2.2776897,
similarity=38.140453]
22: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.6311253,
final_local_scores=0.5875725, pem=-1.9963025,
similarity=32.48025]
23: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.733089,
final_local_scores=0.6952951, pem=-3.4989398,
similarity=43.866486]
24: 928267 Ian_Healy
    [pem_pos=0, final_scores=0.8781683,
final_local_scores=0.93374175, pem=0.0, similarity=31.954428]
25: 26750 Sri_Lanka
    [pem_pos=0, final_scores=0.4325406,
final_local_scores=0.4542989, pem=-0.3027737,
similarity=19.293428]
```

Sample: b'167_PRESS'
PRESS DIGEST - [1 1 ANGOLA 1] - AUG 23 . [2 2 LUANDA 2]
1996-08-23 These are the leading stories in the [3 3 Angolan 3
] press on Friday . [4 4 Reuters 4] has not verified these
stories and does not vouch for their accuracy . - - - -
JORNAL DE ANGOLA - The [5 5 Angolan 5] Chief of State
addressed a letter to [6 6 UN Security Council 6] proposing
dates for the conclusion of the peace process in [7 7 Angola 7

] . He proposed definite dates , August 25 for return of [8 8
Unita 8] generals to the joint army , September 5 for the
beginning of the formation of the Government of National
Unity and Reconciliation . Until this date the free
circulation of peoples and goods should be guaranteed , the
government administration installed in all areas and the [9 9
Unita 9] deputies should occupy their places in the 1 National
Assembly . The president justified his proposal by the delays
verified in the peace process , including the fact that areas
under [10 10 Unita 10] control or occupation have not been
effectively demilitarised , where the [11 11 Unita 11]
military forces have been substituted by their so-called
police . - President Dos Santos proposes the establishment by
[12 12 UN Security Council 12] of definitive and final
timetable for the tasks and obligations under the [13 13 Lusa
ka Agreement 13] and the sending of a mission of SC , as soon
as possible , to supervise the execution of the agreement .

1: 701 Angola
2: 18402 Luanda
3: 701 Angola
4: 18998750 Reuters
5: 701 Angola
6: 31956 United_Nations_Security_Council
7: 701 Angola
8: 245423 UNITA
9: 245423 UNITA
10: 245423 UNITA
11: 245423 UNITA
12: 31956 United_Nations_Security_Council
13: 12271472 Lusaka_Protocol
1: 4014249 National_Assembly_(Angola)
[winner: 140559 National_Assembly, pem_pos=0,
final_scores=-0.39100242, final_local_scores=-0.3878507,
pem=-2.206462, similarity=4.514557]
RECALL MISS: 4014249 National_Assembly_(Angola)]
1: 701 Angola
[pem_pos=0, final_scores=0.52340883,
final_local_scores=0.508355, pem=-0.39249876,
similarity=21.39451]
2: 18402 Luanda
[pem_pos=0, final_scores=0.38577253,
final_local_scores=0.39587831, pem=0.0, similarity=15.903511]
3: 701 Angola
[pem_pos=0, final_scores=0.5174275,
final_local_scores=0.5018213, pem=-0.020324398,
similarity=19.175589]
4: 18998750 Reuters
[pem_pos=0, final_scores=0.2814334,
final_local_scores=0.3466254, pem=0.0, similarity=14.433706]
5: 701 Angola
[pem_pos=0, final_scores=0.6160329,
final_local_scores=0.6095327, pem=-0.020324398,

```
similarity=22.389914]
6: 31956 United_Nations_Security_Council
    [pem_pos=0, final_scores=0.44925582,
final_local_scores=0.45789808, pem=0.0, similarity=17.754305]
7: 701 Angola
    [pem_pos=0, final_scores=0.52587384,
final_local_scores=0.51104766, pem=-0.39249876,
similarity=21.474863]
8: 245423 UNITA
    [pem_pos=0, final_scores=0.5719194,
final_local_scores=0.5447394, pem=0.0, similarity=20.345823]
9: 245423 UNITA
    [pem_pos=0, final_scores=0.56091994,
final_local_scores=0.5327242, pem=0.0, similarity=19.987267]
10: 245423 UNITA
    [pem_pos=0, final_scores=0.56760746,
final_local_scores=0.5400293, pem=0.0, similarity=20.205263]
11: 245423 UNITA
    [pem_pos=0, final_scores=0.5089314,
final_local_scores=0.47593462, pem=0.0, similarity=18.292551]
12: 31956 United_Nations_Security_Council
    [pem_pos=0, final_scores=0.46499932,
final_local_scores=0.47509545, pem=0.0, similarity=18.26751]
13: 12271472 Lusaka_Protocol
    [pem_pos=0, final_scores=0.47742963,
final_local_scores=0.49741393, pem=0.0, similarity=18.933537]
```

Sample: b'168_FORECAST'

FORECAST - S.AFRICAN COMPANY RESULTS CONSENSUS . [1 1 JOHANN
ESBURG 1] 1996-08-23 Analysts estimates of major [2 2 South Af
rican 2] company results expected next week include the
following (all figures cents per share) : DAY--COMPANY----
PERIOD--CONSENSUS----RANGE-----PVS MON Gencor YR EPS 93.12
92.0-94.5 73.8 MON Gencor YR DIV 25.75 25.0-27.0 20.0 MON
Primedia YR EPS N / A 149.1 MON Primedia YR DIV N / A 123.2
MON Distillers YR EPS N / A 71.8 MON Distillers YR DIV N / A
49.0 TUE Iscor YR EPS 29.7 26.0-32.0 38.0 TUE Iscor YR DIV
15.0 14.5-16.5 16.5 TUE McCarthy YR EPS 125.3 112.0-149.0
93.2 TUE McCarthy YR DIV 36.8 32.0-43.0 28.0 WED Imphold YR
EPS 172.7 170.4-175.0 115.1 WED Imphold YR DIV 67.5 66.6-68.4
45.0 THU M&R YR EPS 113.0 112.1-113.4 126.0 THU M&R YR DIV
31.7 10.5-42.3 47.0 THU JD Group YR EPS 143.7 138.0-149.0
111.2 THU JD Group YR DIV 41.8 41.0-42.5 33.0 oo00oo -- [3 3 J
ohannesburg 3] newsroom , +27 11 482 1003

1: 58893 Johannesburg
2: 17416221 South_Africa
3: 58893 Johannesburg

1: 58893 Johannesburg
 [pem_pos=0, final_scores=0.60331404,
final_local_scores=0.55028504, pem=-0.09562059,
similarity=21.031317]
2: 17416221 South_Africa

```
[pem_pos=0, final_scores=0.38346046,
final_local_scores=0.3663026, pem=-0.1389451,
similarity=15.776518]
3: 58893 Johannesburg
    [pem_pos=0, final_scores=0.6020681,
final_local_scores=0.5489241, pem=-0.09562059,
similarity=20.990704]

Sample: b'169_Ulster'
    Ulster Petroleums Ltd Q2 net profit falls . [1 1 CALGARY 1]
1996-08-23 1996 1995 Shr C$ 0.04 C$ 0.08 Net 1,196 2,232 Cash
flow / shr 0.39 0.41 Revs 20,167 18,623 6 MONTHS Shr C$ 0.12
C$ 0.15 Net 3,674 4,271 Cash flow / shr 0.86 0.81 Revs 41,752
35,711 ( All data above 000s except per share numbers ) --
Reuters Toronto Bureau 416 941-8100
1: 15895358 Calgary
1: 15895358 Calgary
    [pem_pos=0, final_scores=0.20790224,
final_local_scores=0.4996888, pem=-0.17577672,
similarity=19.957325]

Sample: b'170_Nigerian'
    [1 1 Nigerian 1] terms jeopardize [2 2 Commonwealth 2] trip-
Canada . [3 3 OTTAWA 3] 1996-08-23 [4 4 Commonwealth 4]
ministers concerned about human rights in [5 5 Nigeria 5] may
cancel a planned trip there because of government
restrictions on their mission , [6 6 Canadian 6] Foreign
Minister [7 7 Lloyd Axworthy 7] said on Friday . " The
reaction of the regime there is such that many of us feel
that the mission under the present circumstances should n't
go ahead , " [8 8 Axworthy 8] said . [9 9 Commonwealth 9]
foreign ministers will meet in [10 10 London 10] on Wednesday
to discuss what to do , he added .
1: 21383 Nigeria
2: 21175158 Commonwealth_of_Nations
3: 22219 Ottawa
4: 21175158 Commonwealth_of_Nations
5: 21383 Nigeria
6: 5042916 Canada
7: 427533 Lloyd_Axworthy
8: 427533 Lloyd_Axworthy
9: 21175158 Commonwealth_of_Nations
10: 17867 London
1: 21383 Nigeria
    [pem_pos=0, final_scores=0.4940428,
final_local_scores=0.5136721, pem=-0.075027294,
similarity=19.826723]
2: 21175158 Commonwealth_of_Nations
    [pem_pos=0, final_scores=0.4973359,
final_local_scores=0.4736132, pem=-0.69133765,
similarity=21.982878]
3: 22219 Ottawa
```

```
[pem_pos=0, final_scores=0.47025,
final_local_scores=0.46408272, pem=-0.34965137,
similarity=19.840326]
4: 21175158 Commonwealth_of_Nations
    [pem_pos=0, final_scores=0.5448387,
final_local_scores=0.5255028, pem=-0.69133765,
similarity=23.531366]
5: 21383 Nigeria
    [pem_pos=0, final_scores=0.5482912,
final_local_scores=0.5729302, pem=-0.31455088,
similarity=22.89767]
6: 5042916 Canada
    [pem_pos=0, final_scores=0.67637366,
final_local_scores=0.6858244, pem=-0.41303828,
similarity=26.802246]
7: 427533 Lloyd_Axworthy
    [pem_pos=0, final_scores=0.8027918,
final_local_scores=0.8888586, pem=0.0, similarity=30.615026]
8: 427533 Lloyd_Axworthy
    [pem_pos=0, final_scores=0.36428648,
final_local_scores=0.40985817, pem=0.0, similarity=16.320698]
9: 21175158 Commonwealth_of_Nations
    [pem_pos=0, final_scores=0.45277393,
final_local_scores=0.424936, pem=-0.69133765,
similarity=20.530252]
10: 17867 London
    [pem_pos=0, final_scores=0.5403266,
final_local_scores=0.554184, pem=-0.09086737,
similarity=21.12182]
```

Sample: b'171_Mid-tier'

Mid-tier golds up in heavy trading . [1 1 TORONTO 1]
1996-08-23 Investors gave into gold fever Friday morning ,
with heavy trading in a handful of [2 2 Toronto-based 2] gold
companies . TVX Gold Inc was up C\$ 0.30 to C\$ 11.55 in
trading of 780,000 shares , while Kinross Gold Corp gained C\$
0.25 to C\$ 11 in volume of 720,000 shares . And Scorpion
Minerals Inc , a junior gold exploration company with five [3
3 Indonesian 3] mining properties , was up C\$ 0.50 to C\$ 6 ,
with about 120,000 shares changing hands . TVX and Kinross
rose after recent buy recommendations from [4 4 U.S. 4]
brokers , analysts said . But Scorpion was raising a lot of
eyebrows after it issued a release Friday morning saying it
was not aware of any developments that could have affected
the stock . The company was formed this year and a couple of
analysts have been on their properties , said one analyst .
Exploration results are expected soon . -- Reuters Toronto
Bureau 416 941-8100

```
1: 64646 Toronto
2: 64646 Toronto
3: 14579 Indonesia
4: 3434750 United_States
```

```
1: 64646 Toronto
    [pem_pos=0, final_scores=0.6171343,
final_local_scores=0.59410095, pem=-0.12291032,
similarity=22.487276]
2: 64646 Toronto
    [pem_pos=0, final_scores=0.5722284,
final_local_scores=0.54504806, pem=0.0, similarity=20.355034]
3: 14579 Indonesia
    [pem_pos=1, final_scores=0.42078072,
final_local_scores=0.47349793, pem=-0.74208045,
similarity=22.255384]
4: 3434750 United_States
    [pem_pos=0, final_scores=0.4802583,
final_local_scores=0.46152514, pem=-0.14075686,
similarity=18.628002]
```

Sample: b'172_RESEARCH'

RESEARCH ALERT - Unitog Co upgraded . - Barrington Research Associates Inc said Friday it upgraded Unitog Co to a near-term outperform from a long-term outperform rating . - Analyst Alexander Paris said he expected consistent 20 percent earnings growth after an estimated gain of 18 percent for 1996 . - The stock closed unchanged at 27 , down from a recent high of 30 . -- [1 1 Chicago 1] newsdesk , 312-408-8787

```
1: 6886 Chicago
```

```
1: 6886 Chicago
```

```
    [pem_pos=0, final_scores=0.44751328,
final_local_scores=0.76142746, pem=-0.20658566,
similarity=27.935673]
```

Sample: b'173_Buffett'

[1 1 Buffett 1] raises Property Capital stake . [2 2 WASHING TON 2] 1996-08-23 [3 3 Omaha 3] billionaire [4 4 Warren Buffet 4] said Friday he raised his stake in Property Capital Trust to 8.0 percent from 6.7 percent . In a filing with the Securities and Exchange Commission , [5 5 Buffett 5] said he bought 62,900 additional common shares of the [6 6 Boston-based 6] real estate investment trust at prices ranging from \$ 7.65 to \$ 8.02 a share . The purchases increased his holding in the company to 725,900 shares , which was purchased for a total of \$ 6.2 million , he said . [7 7 Buffet 7] , who is well-known as a long-term investor , is chairman of [8 8 Berkshire Hathaway Inc 8] , a holding company through which he holds investments in several large [9 9 U.S. 9] companies .

```
1: 211518 Warren_Buffett
```

```
2: 108956 Washington,_D.C.
```

```
3: 46159 Omaha,_Nebraska
```

```
4: 211518 Warren_Buffett
```

```
5: 211518 Warren_Buffett
```

```
6: 24437894 Boston
```

```
7: 211518 Warren_Buffett
```

```
8: 314333 Berkshire_Hathaway
9: 3434750 United_States
1: 211518 Warren_Buffett
    [pem_pos=0, final_scores=0.5462664,
final_local_scores=0.500594, pem=0.0, similarity=19.028437]
2: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.53500193,
final_local_scores=0.5844107, pem=-1.7496481,
similarity=31.044552]
3: 46159 Omaha,_Nebraska
    [pem_pos=0, final_scores=0.46915174,
final_local_scores=0.5168365, pem=-0.32942623,
similarity=21.304617]
4: 211518 Warren_Buffett
    [pem_pos=0, final_scores=0.5468008,
final_local_scores=0.49498755, pem=0.0, similarity=18.861128]
5: 211518 Warren_Buffett
    [pem_pos=0, final_scores=0.55859333,
final_local_scores=0.5140593, pem=0.0, similarity=19.430267]
6: 24437894 Boston
    [pem_pos=0, final_scores=0.44206053,
final_local_scores=0.4786358, pem=0.0, similarity=18.37316]
7: 211518 Warren_Buffett
    [pem_pos=0, final_scores=0.57120657,
final_local_scores=0.5278374, pem=0.0, similarity=19.841434]
8: 314333 Berkshire_Hathaway
    [pem_pos=0, final_scores=0.48551565,
final_local_scores=0.48579997, pem=0.0, similarity=18.586952]
9: 3434750 United_States
    [pem_pos=0, final_scores=0.4373364,
final_local_scores=0.4784171, pem=-0.14075686,
similarity=19.13209]
```

Sample: b'174_Colombia'

```
[1 1 Colombia 1] , [2 2 U.S. 2] reach aviation agreement . [3
3 MIAMI 3] 1996-08-23 The [4 4 U.S. 4] and [5 5 Colombian 5]
governments reached an agreement that will allow [6 6 AMR Corp
6] 's [7 7 American Airlines 7] to operate three round-trip
flights between [8 8 New York 8] and [9 9 Bogota 9] , the [10
10 Department of Transportation 10] said Friday . Under the
agreement , which followed talks in [11 11 Miami 11] this
week , [12 12 AMR 12] also will be allowed to shift up to
four of the weekly flights it now operates between [13 13 Mia
mi 13] and [14 14 Colombia 14] to its [15 15 New York 15]
gateway . The [16 16 United States 16] also will be able to
designate one new all-cargo carrier for service between the
two nations after two years . [17 17 Colombia 17] was
permitted to add a single additional round-trip flight to its
current [18 18 New York 18] service , although it will not be
able to do so while under Category Two ( Conditional ) status
under the [19 19 Federal Aviation Administration 19] 's
International Aviation Safety program . [20 20 Colombia 20]
```

would be allowed to add new service when its safety assessment has been improved , the department said . With the exception of the new services just agreed to , the governments of the two nations have agreed to maintain their current level of routes and airlines for the next 2-1/2 years , the agreement said . The agreement resolved a dispute that arose in June when [21 21 Colombia 21] turned down [22 22 American 22] 's request to operate flights between [23 23 New York 23] and [24 24 Bogota 24] , a denial that prompted the [2 5 25 United States 25] to charge that the [26 26 Colombians 26] were breaking a bilateral aviation agreement and to propose sanctions against one of two [27 27 Colombian 27] airlines , [28 28 Avianca 28] and [29 29 ACES 29] .

1: 5222 Colombia
2: 3434750 United_States
3: 53846 Miami
4: 3434750 United_States
5: 5222 Colombia
6: 1042296 AMR_Corporation
7: 2386 American_Airlines
8: 645042 New_York_City
9: 211271 Bogotá
10: 4162678 New_York_City_Department_of_Transportation
11: 53846 Miami
12: 1042296 AMR_Corporation
13: 53846 Miami
14: 5222 Colombia
15: 645042 New_York_City
16: 3434750 United_States
17: 5222 Colombia
18: 645042 New_York_City
19: 11186 Federal_Aviation_Administration
20: 5222 Colombia
21: 5222 Colombia
22: 3434750 United_States
23: 645042 New_York_City
24: 211271 Bogotá
25: 3434750 United_States
26: 5222 Colombia
27: 5222 Colombia
28: 82195 Avianca
29: 487878 ACES_Colombia
1: 5222 Colombia
[pem_pos=0, final_scores=0.5920076,
final_local_scores=0.57960683, pem=-0.22328302,
similarity=22.600584]
2: 3434750 United_States
[pem_pos=0, final_scores=0.4428392,
final_local_scores=0.40128142, pem=-0.14075686,
similarity=16.83021]
3: 53846 Miami
[pem_pos=0, final_scores=0.50870234,

```
final_local_scores=0.48179597, pem=-0.35396606,
similarity=20.392387]
4: 3434750 United_States
    [pem_pos=0, final_scores=0.44599015,
final_local_scores=0.40472335, pem=-0.14075686,
similarity=16.932922]
5: 5222 Colombia
    [pem_pos=0, final_scores=0.5546734,
final_local_scores=0.5388248, pem=-0.32036036,
similarity=21.911488]
6: 1042296 AMR_Corporation
    [pem_pos=0, final_scores=0.52503127,
final_local_scores=0.61534894, pem=0.0, similarity=22.452953]
7: 2386 American_Airlines
    [pem_pos=0, final_scores=0.44413465,
final_local_scores=0.4716723, pem=0.0, similarity=18.165356]
8: 645042 New_York_City
    [pem_pos=1, final_scores=0.49084824,
final_local_scores=0.46545422, pem=-1.7906687,
similarity=27.71773]
9: 211271 Bogotá
    [pem_pos=0, final_scores=0.46615314,
final_local_scores=0.44445723, pem=0.0, similarity=17.353203]
10: 4162678 New_York_City_Department_of_Transportation
    [pem_pos=6, final_scores=0.31022215,
final_local_scores=0.42512077, pem=-3.0454955,
similarity=33.338043]
11: 53846 Miami
    [pem_pos=0, final_scores=0.5506486,
final_local_scores=0.52761596, pem=-0.35396606,
similarity=21.759747]
12: 1042296 AMR_Corporation
    [pem_pos=3, final_scores=0.3892368,
final_local_scores=0.46701413, pem=-3.2168412,
similarity=35.52003]
13: 53846 Miami
    [pem_pos=0, final_scores=0.4826848,
final_local_scores=0.45337582, pem=-0.35396606,
similarity=19.544273]
14: 5222 Colombia
    [pem_pos=0, final_scores=0.45933598,
final_local_scores=0.4346832, pem=-0.22328302,
similarity=18.275774]
15: 645042 New_York_City
    [pem_pos=1, final_scores=0.47717804,
final_local_scores=0.45052165, pem=-1.7906687,
similarity=27.272112]
16: 3434750 United_States
    [pem_pos=0, final_scores=0.48688954,
final_local_scores=0.4493997, pem=-0.073457584,
similarity=17.900171]
17: 5222 Colombia
```

```
[pem_pos=0, final_scores=0.5046908,
final_local_scores=0.48422652, pem=-0.22328302,
similarity=19.754246]
18: 645042 New_York_City
    [pem_pos=1, final_scores=0.47279513,
final_local_scores=0.44573396, pem=-1.7906687,
similarity=27.129238]
19: 11186 Federal_Aviation_Administration
    [pem_pos=0, final_scores=0.47239506,
final_local_scores=0.5407144, pem=0.0, similarity=20.225708]
20: 5222 Colombia
    [pem_pos=0, final_scores=0.6104438,
final_local_scores=0.5997455, pem=-0.22328302,
similarity=23.201563]
21: 5222 Colombia
    [pem_pos=0, final_scores=0.6263488,
final_local_scores=0.6171193, pem=-0.22328302,
similarity=23.720032]
22: 3434750 United_States
    [pem_pos=0, final_scores=0.42384952,
final_local_scores=0.3805381, pem=-0.12994486,
similarity=16.15239]
23: 645042 New_York_City
    [pem_pos=1, final_scores=0.45657295,
final_local_scores=0.42801368, pem=-1.7906687,
similarity=26.60043]
24: 211271 Bogotá
    [pem_pos=0, final_scores=0.47232264,
final_local_scores=0.45119643, pem=0.0, similarity=17.554314]
25: 3434750 United_States
    [pem_pos=0, final_scores=0.50275195,
final_local_scores=0.46672702, pem=-0.073457584,
similarity=18.417252]
26: 5222 Colombia
    [pem_pos=0, final_scores=0.5688744,
final_local_scores=0.5543374, pem=-0.22300437,
similarity=21.844978]
27: 5222 Colombia
    [pem_pos=0, final_scores=0.5020252,
final_local_scores=0.48131466, pem=-0.32036036,
similarity=20.19527]
28: 82195 Avianca
    [pem_pos=0, final_scores=0.2378849,
final_local_scores=0.2845563, pem=0.0, similarity=12.581441]
29: 487878 ACES_Colombia
    [pem_pos=9, final_scores=0.44981462,
final_local_scores=0.59743917, pem=-2.7725887,
similarity=36.996258]
```

Sample: b'175_Clean'

Clean tanker fixtures and enquiries - 1754 [1 1 GMT 1] . [2 2 LONDON 2] 1996-08-23 LATEST FIXTURES - MIDEAST GULF / [3 3 RED

SEA 3] Konpolis 75 1/9 [4 4 Mideast 4] / [5 5 Indonesia 5]
W112.5 KPC . TBN 30 6/9 [6 6 Mideast 6] / W.C. India W200 ,
E.C.India W195 [7 7 IOC 7] . - [8 8 ASIA PACIFIC 8] Petrobulk
Rainbow 28 24/8 [9 9 Okinawa 9] / [10 10 Inchon 10] \$ 190,000
Honam . - [11 11 MED 11] / [12 12 BLACK SEA 12] TBN 30 15/9 [1
3 13 Constanza 13] / [14 14 Inia 14] \$ 700,000 [15 15 IOC 15]
. - UK / CONT Port Christine 36,5 3/9 [16 16 Pembroke 16] /
US W145 Stentex . - [17 17 WESTERN HEMISPHERE 17] Kpaitan
Stankov 69 31/8 [18 18 St Croix 18] / USAC W125 Hess . AP
Moller 30 31/8 Caribs / [19 19 Japan 19] \$ 875,000 [20 20 BP 2
0] . Tiber 29 2/9 Caribs / options W265 Stinnes .

- MIDDAY FIXTURES - MIDEAST GULF / [21 21 RED SEA 21]
Tenacity 70 24/08 [22 22 Mideast 22] / [23 23 South Korea 23]
W145 [24 24 Samsung 24] . SKS Tana 70 03/09 [25 25 Mideast 25]
/ [26 26 Japan 26] W145 CNR . Northsea Chaser 55 12/09 [27 27
Mideast 27] / [28 28 Japan 28] W167.5 Jomo . Sibonina 55
13/09 [29 29 Red Sea 29] / [30 30 Japan 30] W160 [31 31 Marub
eni 31] . - [32 32 ASIA 32] / PACIFIC Neptune Crux 30 02/09 [3
3 33 Singapore 33] / options \$ 185,000 Sietco . World Bridge
30 03/09 [34 34 South Korea 34] / [35 35 Japan 35] rnr CNR .
Fulmar 30 28/08 [36 36 Ulsan 36] / [37 37 Yosu 37] \$ 105,000
LG Caltex . - [1 MED 1] / [38 38 BLACK SEA 38] Hemina 33
05/09 [39 39 Eleusis 39] / UKCM W155 CNR . -- London
Newsroom , +44 171 542 8980

1: 12701 Greenwich_Mean_Time
2: 17867 London
3: 26191 Red_Sea
4: 19323 Middle_East
5: 14579 Indonesia
6: 19323 Middle_East
7: 7971634 Indian_Oil_Corporation
8: 318083 Pacific_Rim
9: 22477 Okinawa_Prefecture
10: 68074 Incheon
11: 19006 Mediterranean_Sea
12: 3386 Black_Sea
13: 265716 Constanța
14: 23298531 Inia,_Paphos
15: 7971634 Indian_Oil_Corporation
16: 116906 Pembroke,_Massachusetts
17: 73162 Western_Hemisphere
18: 229431 Saint_Croix,_U.S._Virgin_Islands
19: 15573 Japan
20: 18998720 BP
21: 26191 Red_Sea
22: 19323 Middle_East
23: 27019 South_Korea
24: 9177773 Samsung_Heavy_Industries
25: 19323 Middle_East
26: 15573 Japan
27: 19323 Middle_East

```
28: 15573 Japan
29: 26191 Red_Sea
30: 15573 Japan
31: 1132575 Marubeni
32: 689 Asia
33: 27318 Singapore
34: 27019 South_Korea
35: 15573 Japan
36: 323111 Ulsan
37: 386393 Yeosu
38: 3386 Black_Sea
39: 80628 Eleusis
1: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.631299,
final_local_scores=0.7185388, pem=0.0, similarity=25.532345]
2: 17867 London
    [pem_pos=0, final_scores=0.68453383,
final_local_scores=0.7215809, pem=-0.09086737,
similarity=26.117277]
3: 26191 Red_Sea
    [pem_pos=0, final_scores=0.4590538,
final_local_scores=0.46895224, pem=0.0, similarity=18.084183]
4: 19323 Middle_East
    [pem_pos=0, final_scores=0.624096,
final_local_scores=0.6074932, pem=0.0, similarity=22.218523]
5: 14579 Indonesia
    [pem_pos=0, final_scores=0.73898065,
final_local_scores=0.75206274, pem=-0.23151328,
similarity=27.79177]
6: 19323 Middle_East
    [pem_pos=0, final_scores=0.5797244,
final_local_scores=0.559024, pem=0.0, similarity=20.772104]
7: 7971634 Indian_Oil_Corporation
    [pem_pos=2, final_scores=0.6212192,
final_local_scores=0.7517361, pem=-1.36746,
similarity=33.95948]
8: 318083 Pacific_Rim
    [pem_pos=1, final_scores=0.23712729,
final_local_scores=0.24013186, pem=-1.3348545,
similarity=18.514872]
9: 22477 Okinawa_Prefecture
    [pem_pos=0, final_scores=0.31128567,
final_local_scores=0.3060672, pem=-0.15917227,
similarity=14.088972]
10: 68074 Incheon
    [pem_pos=0, final_scores=0.3923481,
final_local_scores=0.41721743, pem=-0.4443909,
similarity=18.95698]
11: 19006 Mediterranean_Sea
    [pem_pos=12, final_scores=0.3241375,
final_local_scores=0.29394096, pem=-2.8332133,
similarity=28.268951]
```

```
12: 3386 Black_Sea
    [pem_pos=0, final_scores=0.5019961,
final_local_scores=0.5111608, pem=0.0, similarity=19.343771]
13: 265716 Constanța
    [pem_pos=2, final_scores=0.44942605,
final_local_scores=0.50751156, pem=-1.4556266,
similarity=27.150793]
14: 23298531 Inia,_Paphos
    [pem_pos=1, final_scores=0.39404368,
final_local_scores=0.6228243, pem=-0.9385614,
similarity=27.780075]
15: 7971634 Indian_Oil_Corporation
    [pem_pos=2, final_scores=0.59915715,
final_local_scores=0.72763675, pem=-1.36746,
similarity=33.240307]
16: 116906 Pembroke,_Massachusetts
    [pem_pos=1, final_scores=0.28152013,
final_local_scores=0.37692624, pem=-2.0949647,
similarity=26.730688]
17: 73162 Western_Hemisphere
    [pem_pos=0, final_scores=0.41758662,
final_local_scores=0.4130546, pem=0.0, similarity=16.416086]
18: 229431 Saint_Croix,_U.S._Virgin_Islands
    [pem_pos=0, final_scores=0.32703066,
final_local_scores=0.37056428, pem=0.0, similarity=15.148089]
19: 15573 Japan
    [pem_pos=0, final_scores=0.76313835,
final_local_scores=0.7647071, pem=-0.13549899,
similarity=27.646961]
20: 18998720 BP
    [pem_pos=0, final_scores=0.5358084,
final_local_scores=0.6079719, pem=0.0, similarity=22.232807]
21: 26191 Red_Sea
    [pem_pos=0, final_scores=0.446904,
final_local_scores=0.4556805, pem=0.0, similarity=17.688128]
22: 19323 Middle_East
    [pem_pos=0, final_scores=0.6465407,
final_local_scores=0.6320107, pem=0.0, similarity=22.950172]
23: 27019 South_Korea
    [pem_pos=0, final_scores=0.49550146,
final_local_scores=0.5011764, pem=-0.22449642,
similarity=20.266665]
24: 9177773 Samsung_Heavy_Industries
    [pem_pos=3, final_scores=0.35915107,
final_local_scores=0.4636069, pem=-1.3862944,
similarity=25.463552]
25: 19323 Middle_East
    [pem_pos=0, final_scores=0.64510614,
final_local_scores=0.63044363, pem=0.0, similarity=22.903408]
26: 15573 Japan
    [pem_pos=0, final_scores=0.8145233,
final_local_scores=0.8208373, pem=-0.13549899,
```

```
similarity=29.322]
27: 19323 Middle_East
    [pem_pos=0, final_scores=0.61822134,
final_local_scores=0.60107607, pem=0.0, similarity=22.027021]
28: 15573 Japan
    [pem_pos=0, final_scores=0.81798655,
final_local_scores=0.82462037, pem=-0.13549899,
similarity=29.434893]
29: 26191 Red_Sea
    [pem_pos=0, final_scores=0.371805,
final_local_scores=0.37364626, pem=0.0, similarity=15.240062]
30: 15573 Japan
    [pem_pos=0, final_scores=0.8189164,
final_local_scores=0.825636, pem=-0.13549899,
similarity=29.465202]
31: 1132575 Marubeni
    [pem_pos=0, final_scores=0.43542427,
final_local_scores=0.5676838, pem=0.0, similarity=21.03053]
32: 689 Asia
    [pem_pos=0, final_scores=0.63146085,
final_local_scores=0.61754805, pem=-0.1494309,
similarity=23.331207]
33: 27318 Singapore
    [pem_pos=0, final_scores=0.5195573,
final_local_scores=0.5074226, pem=-0.1922004,
similarity=20.277431]
34: 27019 South_Korea
    [pem_pos=0, final_scores=0.40263218,
final_local_scores=0.3997308, pem=-0.22449642,
similarity=17.239325]
35: 15573 Japan
    [pem_pos=0, final_scores=0.61112696,
final_local_scores=0.59865767, pem=-0.13549899,
similarity=22.691715]
36: 323111 Ulsan
    [pem_pos=0, final_scores=0.51643616,
final_local_scores=0.56561935, pem=0.0, similarity=20.968922]
37: 386393 Yeosu
    [pem_pos=1, final_scores=0.4420284,
final_local_scores=0.5041225, pem=-0.36593428,
similarity=21.123741]
38: 3386 Black_Sea
    [pem_pos=0, final_scores=0.48746324,
final_local_scores=0.49528587, pem=0.0, similarity=18.870031]
39: 80628 Eleusis
    [pem_pos=0, final_scores=0.52013206,
final_local_scores=0.61186785, pem=0.0, similarity=22.34907]
1: 19006 Mediterranean_Sea
    [pem_pos=12, final_scores=0.22397415,
final_local_scores=0.1845277, pem=-2.8332133,
similarity=25.003841]
```

Sample: b'176_CRICKET'

CRICKET - [1 1 PAKISTAN 1] 229-1 V [2 2 ENGLAND 2] - close .
[3 3 Saeed Anwar 3] not out 116 Aamir Sohail c [4 4 Cork 4] b
[5 5 Croft 5] 46 [6 6 Ijaz Ahmed 6] not out 58 Extras 9 Fall
of wicket - 1-106 To bat - [7 7 Inzamam-ul-Haq 7] , [8 8 Salim
Malik 8] , [9 9 Asif Mujtaba 9] , [10 10 Wasim Akram 10] , [11
11 Moin Khan 11] , [12 12 Mushtaq Ahmed 12] , [13 13 Waqar Yo
unis 13] , Mohammad Akam [14 14 England 14] 326 all out

1: 759295 Pakistan_national_cricket_team
2: 759125 England_cricket_team
3: 894437 Saeed_Anwar
4: 2036840 Dominic_Cork
5: 1160680 Robert_Croft
6: 938863 Ijaz_Ahmed_(cricketer)
7: 896403 Inzamam-ul-Haq
8: 938861 Saleem_Malik
9: 3976606 Asif_Mujtaba
10: 501393 Wasim_Akram
11: 554132 Moin_Khan
12: 896405 Mushtaq_Ahmed
13: 894433 Waqar_Younis
14: 759125 England_cricket_team

1: 759295 Pakistan_national_cricket_team
[pem_pos=1, final_scores=0.78806955,
final_local_scores=0.7742242, pem=-2.7556365,
similarity=42.179688]

2: 759125 England_cricket_team
[pem_pos=2, final_scores=0.6895594,
final_local_scores=0.6681327, pem=-3.3054674,
similarity=42.003773]

3: 894437 Saeed_Anwar
[pem_pos=0, final_scores=0.8190357,
final_local_scores=0.82699835, pem=0.0, similarity=28.768993]

4: 2036840 Dominic_Cork
[pem_pos=8, final_scores=1.13764,
final_local_scores=1.1845918, pem=-3.606134,
similarity=59.05101]

5: 1160680 Robert_Croft
[pem_pos=7, final_scores=0.9679844,
final_local_scores=1.0291367, pem=-3.16984,
similarity=52.03929]

6: 938863 Ijaz_Ahmed_(cricketer)
[pem_pos=0, final_scores=0.6271631,
final_local_scores=0.6209922, pem=0.0, similarity=22.621359]

7: 896403 Inzamam-ul-Haq
[pem_pos=0, final_scores=0.8602647,
final_local_scores=0.8632036, pem=0.0, similarity=29.849428]

8: 938861 Saleem_Malik
[pem_pos=0, final_scores=0.72501713,
final_local_scores=0.7489689, pem=0.0, similarity=26.44044]

9: 3976606 Asif_Mujtaba
[pem_pos=0, final_scores=0.71830577,

```
final_local_scores=0.75304437, pem=0.0, similarity=26.56206]
10: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.74670416,
final_local_scores=0.7409494, pem=0.0, similarity=26.20112]
11: 554132 Moin_Khan
    [pem_pos=0, final_scores=0.72073334,
final_local_scores=0.75814116, pem=0.0, similarity=26.714157]
12: 896405 Mushtaq_Ahmed
    [pem_pos=0, final_scores=0.6027123,
final_local_scores=0.60069585, pem=0.0, similarity=22.015675]
13: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.7244006,
final_local_scores=0.7229196, pem=0.0, similarity=25.663074]
14: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.27742946,
final_local_scores=0.22508222, pem=-3.3054674,
similarity=28.78226]
```

Sample: b'177_RTRS'

RTRS - Golf : [1 1 Norman 1] sacks his coach after disappointing season . [2 2 SYDNEY 2] 1996-08-23 World number one golfer [3 3 Greg Norman 3] has sacked his coach [4 4 Butch Harmon 4] after a disappointing season . " [5 5 Butch 5] and I are finished , " [6 6 Norman 6] told reporters on Thursday before the start of the [7 7 World Series of Golf 7] in [8 8 Akron 8] , [9 9 Ohio 9] . [10 10 Norman 10] , a two-time [11 11 British Open 11] champion , parted ways with his long-time mentor after drawing a blank in this year 's four majors , winning two tournaments worldwide . The blonde Australian opened with a level par round of 70 in [12 12 Akron 12] , leaving him four shots adrift of the leaders , [13 13 Americans 13] [14 14 Billy Mayfair 14] and [15 15 Paul Goydos 15] and [16 16 Japan 16] 's Hidemichi Tanaki . On Wednesday [17 17 Norman 17] described this year as his worst on the professional circuit since 1991 , when he failed to win a tournament . " My application this year has been strange , " [18 18 Norman 18] said . " Maybe I have n't been as keyed up as I should have been . " " Sometimes you do n't have it in your head to play . Maybe this was one of those years where I was there , but I was n't 100 percent there , and you have to be 100 percent to perform , " he said . -- Sydney Newsroom

61-2 9373-1800

```
1: 157447 Greg_Norman
2: 27862 Sydney
3: 157447 Greg_Norman
4: 2363110 Butch_Harmon
5: 2363110 Butch_Harmon
6: 157447 Greg_Norman
7: 1536678 World_Series_of_Golf
8: 129842 Akron,_Ohio
9: 22199 Ohio
10: 157447 Greg_Norman
```

```
11: 176725 The_Open_Championship
12: 129842 Akron,_Ohio
13: 3434750 United_States
14: 2080225 Billy_Mayfair
15: 6280981 Paul_Goydos
16: 15573 Japan
17: 157447 Greg_Norman
18: 157447 Greg_Norman
1: 157447 Greg_Norman
    [pem_pos=0, final_scores=0.7529767,
final_local_scores=0.7438195, pem=0.0, similarity=26.28677]
2: 27862 Sydney
    [pem_pos=0, final_scores=0.5478788,
final_local_scores=0.60071003, pem=-0.1443489,
similarity=22.80109]
3: 157447 Greg_Norman
    [pem_pos=0, final_scores=0.6388715,
final_local_scores=0.6191768, pem=0.0, similarity=22.567184]
4: 2363110 Butch_Harmon
    [pem_pos=0, final_scores=0.6758235,
final_local_scores=0.7093906, pem=0.0, similarity=25.259342]
5: 2363110 Butch_Harmon
    [pem_pos=0, final_scores=1.3090152,
final_local_scores=1.4138396, pem=0.0, similarity=46.281506]
6: 157447 Greg_Norman
    [pem_pos=0, final_scores=0.67614675,
final_local_scores=0.65989435, pem=0.0, similarity=23.782276]
7: 1536678 World_Series_of_Golf
    [pem_pos=0, final_scores=0.33112854,
final_local_scores=0.3867066, pem=0.0, similarity=15.629808]
8: 129842 Akron,_Ohio
    [pem_pos=1, final_scores=0.43089038,
final_local_scores=0.4630589, pem=-0.6020413,
similarity=21.182308]
9: 22199 Ohio
    [pem_pos=0, final_scores=0.5667092,
final_local_scores=0.6419248, pem=-0.19147073,
similarity=24.287277]
10: 157447 Greg_Norman
    [pem_pos=0, final_scores=0.63820034,
final_local_scores=0.61844367, pem=0.0, similarity=22.545305]
11: 176725 The_Open_Championship
    [pem_pos=0, final_scores=0.21906517,
final_local_scores=0.19321683, pem=-0.13592334,
similarity=10.594861]
12: 129842 Akron,_Ohio
    [pem_pos=1, final_scores=0.40564,
final_local_scores=0.43547672, pem=-0.6020413,
similarity=20.359201]
13: 3434750 United_States
    [pem_pos=0, final_scores=0.46472245,
final_local_scores=0.5178181, pem=-0.5581965,
```

```
similarity=22.577997]
14: 2080225 Billy_Mayfair
    [pem_pos=0, final_scores=0.7171269,
final_local_scores=0.7929302, pem=0.0, similarity=27.75233]
15: 6280981 Paul_Goydos
    [pem_pos=0, final_scores=0.7118771,
final_local_scores=0.79031336, pem=0.0, similarity=27.67424]
16: 15573 Japan
    [pem_pos=0, final_scores=0.7552319,
final_local_scores=0.8771768, pem=-0.13549899,
similarity=31.003284]
17: 157447 Greg_Norman
    [pem_pos=0, final_scores=0.68383,
final_local_scores=0.66828716, pem=0.0, similarity=24.032734]
18: 157447 Greg_Norman
    [pem_pos=0, final_scores=0.69099295,
final_local_scores=0.6761116, pem=0.0, similarity=24.266232]
```

Sample: b'178_Soccer'

Soccer - [1 1 Arab 1] team breaks new ground in [2 2 Israel 2] . Ori Lewis TAIBE , [3 3 Israel 3] 1996-08-23 For the first time in [4 4 Israeli 4] history , an [5 5 Arab 5] team will take the field when the National League soccer season starts on Saturday . [6 6 Hapoel Taibe 6] fields four [7 7 Jewish 7] players and two foreign imports -- a [8 8 Pole 8] and a [9 9 R omanian 9] . The rest of the side is made up mainly of Moslem Arabs . The club , founded in 1961 , has a loyal following in [10 10 Taibe 10] , an [11 11 Arab 11] town of 28,000 in the heart of [12 12 Israel 12] . But away from their home ground , they face unfriendly crowds who taunt the players with racist abuse . " The very first thing we thought about after we knew we would be promoted was the game against [13 13 Betar Jerusalem 13] , " said Taibe supporter Karem Haj Yihye . Two weeks ago Taibe , coached by [14 14 Pole 14] Wojtek Lazarek , met Betar , a club closely associated with the right-wing [15 15 Likud 15] party , for the first time in a Cup match in [16 16 Jerusalem 16] . Chants from the crowd of " Death to the [17 17 Arabs 17] " , and bottle-throwing during the game marred the match which ended in a goalless draw . One Taibe supporter required hospital treatment for cuts and bruises after a stone struck his head as he was driving from the stadium . " We 're used to hearing the taunts of " Death to the [18 18 Arabs 18] ' , " said Sameh Haj Yihye , a [19 19 Taibe 19] resident who studies at [20 20 Jerusalem 20] 's [21 21 Hebrew University 21] . " But we know that these are only words , nobody has died from hearing them and it only makes us support our team more vehemently . " The dusty town of [22 22 Taibe 22] lacks the amenities of [23 23 J ewish 23] communities and many [24 24 Israeli Arabs 24] have long complained of state discrimination . " There are no parks or empty areas of land around here , so when we want to play a friendly game of soccer we all load up in the car and

travel to [25 25 Tel Aviv 25] , " 60 km (36 miles) away , Sameh Haj Yihye said . The town 's ramshackle 2,500-seat ground is accessible only by two dirt tracks . " We plan to build a 10,000-seat stadium , but it may well be situated elsewhere , " said club chairman Abdul Rahman Haj Yihye . " We will discuss this with the mayor and hopefully a new or refurbished ground will be completed by the start of the new year . " In the meantime , Taibe will play all their heavily policed home matches at the [26 26 Jewish 26] coastal town of Netanya . " We are [27 27 Israelis 27] , there is no question about that , " said Karem Haj Yihye , a hotel waiter . " We do n't have any connection with the [28 28 Palestinians 28] , they live over there , " he said , pointing to the West Bank seven km (four miles) to the east . " We do n't feel our club represents Palestinian Arabs , " said club chairman Abdul Rahman . " We are trying to do all we can to run a professional outfit , we are pleased at any support we get , but do not go out looking to represent the whole [29 29 Arab 2 9] world . "

1: 159433 Arab_world
2: 9282173 Israel
3: 9282173 Israel
4: 9282173 Israel
5: 159433 Arab_world
6: 19486919 Hapoel_Tayibe_F.C.
7: 25955086 Jews
8: 275297 Poles
9: 322831 Romanians
10: 17929012 Taibe,_Galilee
11: 159433 Arab_world
12: 9282173 Israel
13: 1755537 Beitar_Jerusalem_F.C.
14: 275297 Poles
15: 206984 Likud
16: 16043 Jerusalem
17: 2185 Arab_people
18: 2185 Arab_people
19: 17929012 Taibe,_Galilee
20: 16043 Jerusalem
21: 229168 Hebrew_University_of_Jerusalem
22: 17929012 Taibe,_Galilee
23: 25955086 Jews
24: 492331 Arab_citizens_of_Israel
25: 31453 Tel_Aviv
26: 25955086 Jews
27: 181296 Israelis
28: 241405 State_of_Palestine
29: 159433 Arab_world
1: 159433 Arab_world
[pem_pos=14, final_scores=0.51881796, final_local_scores=0.5074005, pem=-3.8052368, similarity=39.925026]

```
2: 9282173 Israel
    [pem_pos=0, final_scores=0.6201548,
final_local_scores=0.58058614, pem=-0.21279564,
similarity=22.572777]
3: 9282173 Israel
    [pem_pos=0, final_scores=0.57168067,
final_local_scores=0.5276355, pem=-0.21279564,
similarity=20.992624]
4: 9282173 Israel
    [pem_pos=0, final_scores=0.49739963,
final_local_scores=0.44649476, pem=-0.34790626,
similarity=19.305975]
5: 159433 Arab_world
    [pem_pos=14, final_scores=0.55953914,
final_local_scores=0.55188227, pem=-3.8052368,
similarity=41.25245]
6: 19486919 Hapoel_Tayibe_F.C.
    [pem_pos=0, final_scores=0.26921743,
final_local_scores=0.40107298, pem=0.0, similarity=16.05853]
7: 25955086 Jews
    [pem_pos=1, final_scores=0.39829057,
final_local_scores=0.35540265, pem=-0.80476934,
similarity=19.072098]
8: 275297 Poles
    [pem_pos=1, final_scores=0.455478,
final_local_scores=0.47682953, pem=-1.8542098,
similarity=28.402739]
9: 322831 Romanians
    [pem_pos=2, final_scores=0.161974,
final_local_scores=0.16023216, pem=-1.8371289,
similarity=18.861952]
10: 17929012 Taibe,_Galilee
    [pem_pos=0, final_scores=0.4092642,
final_local_scores=0.5087094, pem=0.0, similarity=19.270615]
11: 159433 Arab_world
    [pem_pos=14, final_scores=0.5131057,
final_local_scores=0.50116074, pem=-3.8052368,
similarity=39.73882]
12: 9282173 Israel
    [pem_pos=0, final_scores=0.5856703,
final_local_scores=0.5429171, pem=-0.21279564,
similarity=21.448656]
13: 1755537 Beitar_Jerusalem_F.C.
    [pem_pos=0, final_scores=0.33661807,
final_local_scores=0.39052486, pem=0.0, similarity=15.743754]
14: 275297 Poles
    [pem_pos=1, final_scores=0.58663166,
final_local_scores=0.620095, pem=-1.8542098,
similarity=32.678066]
15: 206984 Likud
    [pem_pos=0, final_scores=0.4107557,
final_local_scores=0.44354898, pem=0.0, similarity=17.3261]
```

16: 16043 Jerusalem
[pem_pos=0, final_scores=0.5216936,
final_local_scores=0.49108267, pem=-0.13239029,
similarity=19.464558]
17: 2185 Arab_people
[pem_pos=1, final_scores=0.43778795,
final_local_scores=0.41362768, pem=-1.8745545,
similarity=26.627306]
18: 2185 Arab_people
[pem_pos=1, final_scores=0.43387932,
final_local_scores=0.40935808, pem=-1.8745545,
similarity=26.499893]
19: 17929012 Taibe,_Galilee
[pem_pos=0, final_scores=0.42141664,
final_local_scores=0.52198404, pem=0.0, similarity=19.666758]
20: 16043 Jerusalem
[pem_pos=0, final_scores=0.5164129,
final_local_scores=0.48531437, pem=-0.13239029,
similarity=19.292421]
21: 229168 Hebrew_University_of_Jerusalem
[pem_pos=0, final_scores=0.42979252,
final_local_scores=0.45298815, pem=0.0, similarity=17.607782]
22: 17929012 Taibe,_Galilee
[pem_pos=0, final_scores=0.38521123,
final_local_scores=0.4824351, pem=0.0, similarity=18.486538]
23: 25955086 Jews
[pem_pos=1, final_scores=0.48456097,
final_local_scores=0.44963998, pem=-0.80476934,
similarity=21.884327]
24: 492331 Arab_citizens_of_Israel
[pem_pos=0, final_scores=0.38491434,
final_local_scores=0.3625159, pem=0.0, similarity=14.90791]
25: 31453 Tel_Aviv
[pem_pos=0, final_scores=0.4260921,
final_local_scores=0.40636253, pem=0.0, similarity=16.216381]
26: 25955086 Jews
[pem_pos=1, final_scores=0.4383598,
final_local_scores=0.39917225, pem=-0.80476934,
similarity=20.37827]
27: 181296 Israelis
[pem_pos=0, final_scores=0.28821105,
final_local_scores=0.262451, pem=-0.009432555,
similarity=11.97307]
28: 241405 State_of_Palestine
[pem_pos=5, final_scores=0.37589538,
final_local_scores=0.3937236, pem=-1.7909594,
similarity=25.578726]
29: 159433 Arab_world
[pem_pos=14, final_scores=0.4813246,
final_local_scores=0.46644473, pem=-3.8052368,
similarity=38.702824]

Sample: b'179_Soccer'

Soccer - [1 1 Kennedy 1] and [2 2 Phelan 2] both out of [3 3 Irish 3] squad . [4 4 DUBLIN 4] 1996-08-23 Two players have withdrawn from the [5 5 Republic of Ireland 5] squad for the 1998 [6 6 World Cup 6] qualifying match against [7 7 Liechtenstein 7] on August 31 , the [8 8 Football Association of Ireland 8] said in a statement on Friday . The F.A.I. statement said that [9 9 Liverpool 9] striker [10 10 Mark Kennedy 10] and [11 11 Chelsea 11] defender [12 12 Terry Phelan 12] were both receiving treatment for injuries and would not be travelling to [13 13 Liechtenstein 13] for the game . No replacements had been named . -- [14 14 Damien Lynch 14] , Dublin Newsroom +353 1 6603377

1: 3567921 Mark_Kennedy_(footballer)

2: 5348141 Terry_Phelan

3: 14560 Republic_of_Ireland

4: 8504 Dublin

5: 14560 Republic_of_Ireland

6: 126460 1998_FIFA_World_Cup

7: 652703 Liechtenstein_national_football_team

8: 13373575 Football_Association_of_Ireland

9: 18119 Liverpool_F.C.

10: 3567921 Mark_Kennedy_(footballer)

11: 7473 Chelsea_F.C.

12: 5348141 Terry_Phelan

13: 17810 Liechtenstein

14: 23279581 Damian_Lynch

1: 3567921 Mark_Kennedy_(footballer)
[pem_pos=1, final_scores=0.75701576,
final_local_scores=0.79240435, pem=-0.120628,
similarity=28.392633]

2: 5348141 Terry_Phelan
[pem_pos=0, final_scores=1.0263416,
final_local_scores=1.1051285, pem=0.0, similarity=37.06895]

3: 14560 Republic_of_Ireland
[pem_pos=3, final_scores=0.54618865,
final_local_scores=0.5929695, pem=-2.2576945,
similarity=34.062798]

4: 8504 Dublin
[pem_pos=0, final_scores=0.46643835,
final_local_scores=0.49525994, pem=-0.23150389,
similarity=20.128212]

5: 14560 Republic_of_Ireland
[pem_pos=0, final_scores=0.5333715,
final_local_scores=0.57896864, pem=-0.2122539,
similarity=22.52156]

6: 126460 1998_FIFA_World_Cup
[pem_pos=18, final_scores=0.31334388,
final_local_scores=0.3127042, pem=-4.2172847,
similarity=36.35568]

7: 652703 Liechtenstein_national_football_team
[pem_pos=1, final_scores=0.5265164,

```
final_local_scores=0.55859274, pem=-0.26236424,
similarity=22.186012]
8: 13373575 Football_Association_of_Ireland
    [pem_pos=0, final_scores=0.44210953,
final_local_scores=0.44490695, pem=0.0, similarity=17.366623]
9: 18119 Liverpool_F.C.
    [pem_pos=1, final_scores=0.6269303,
final_local_scores=0.6119154, pem=-1.1702383,
similarity=28.714428]
10: 3567921 Mark_Kennedy_(footballer)
    [pem_pos=1, final_scores=0.50941527,
final_local_scores=0.5219384, pem=-0.120628,
similarity=20.32139]
11: 7473 Chelsea_F.C.
    [pem_pos=0, final_scores=0.6964551,
final_local_scores=0.704905, pem=-0.51707125,
similarity=27.937397]
12: 5348141 Terry_Phelan
    [pem_pos=0, final_scores=0.54366213,
final_local_scores=0.5617726, pem=0.0, similarity=20.854128]
13: 17810 Liechtenstein
    [pem_pos=0, final_scores=0.3799314,
final_local_scores=0.45033097, pem=0.0, similarity=17.528488]
14: 23279581 Damian_Lynch
    [pem_pos=0, final_scores=0.4299156,
final_local_scores=0.51261455, pem=0.0, similarity=19.387154]
```

Sample: b'180_Soccer'

Soccer - [1 1 Manchester United 1] face [2 2 Juventus 2] in [3 3 Europe 3] . [4 4 GENEVA 4] 1996-08-23 [5 5 European 5] champions [6 6 Juventus 6] will face [7 7 English 7] league and cup double winners [8 8 Manchester United 8] in this season 's European Champions ' League . The draw made on Friday pitted [9 9 Juventus 9] , who beat [10 10 Dutch 10] champions [11 11 Ajax Amsterdam 11] 4-2 on penalties in last year 's final , against [12 12 Alex Ferguson 12] 's European hopefuls in group C . The other two teams in the group are last season 's Cup Winners ' Cup runners-up [13 13 Rapid Viena 13] and [14 14 Fenerbahce 14] of Turkey . [15 15 Juventus 15] meet [16 16 United 16] in [17 17 Turin 17] on September 11 , with the return match at [18 18 Old Trafford 18] on November 20 . [19 19 United 19] have dominated the premier league in the 1990s , winning three [20 20 English 20] championships in four years , but have consistently failed in [21 21 Europe 21] , crashing out of the [22 22 European Cup 22] to [23 23 Galatasaray 23] of Turkey and [24 24 Spain 24] 's [25 25 Barcelona 25] at their last two attempts . They have not lifted a European Trophy since 1991 when they beat [26 26 Barcelona 26] in the Cup Winners ' Cup final , and their one and only [27 27 European Cup 27] triumph was way back in 1968 , when they beat [28 28 Benfica 28] of [29 29 Portugal 29] 4-1 at [30 30 Wembley 30] . [31 31 Juventus 31] have won

the [32 32 European Cup 32] twice . Before conquering [33 33 A jax 33] last year they beat [34 34 United 34] 's big [35 35 En glish 35] rivals [36 36 Liverpool 36] in the ill-fated 1985 final in the [37 37 Heysel 37] stadium in [38 38 Brussels 38]

.

1: 19961 Manchester_United_F.C.
2: 16415 Juventus_F.C.
3: 9239 Europe
4: 12521 Geneva
5: 44220 UEFA_Champions_League
6: 16415 Juventus_F.C.
7: 9316 England
8: 19961 Manchester_United_F.C.
9: 16415 Juventus_F.C.
10: 21148 Netherlands
11: 2273 AFC_Ajax
12: 43223 Alex_Ferguson
13: 656932 SK_Rapid_Wien
14: 979542 Fenerbahçe_S.K.
15: 16415 Juventus_F.C.
16: 19961 Manchester_United_F.C.
17: 19450529 Turin
18: 660592 Old_Trafford
19: 19961 Manchester_United_F.C.
20: 9316 England
21: 9239 Europe
22: 44220 UEFA_Champions_League
23: 23608452 Galatasaray_S.K._(football)
24: 26667 Spain
25: 68187 FC_Barcelona
26: 68187 FC_Barcelona
27: 44220 UEFA_Champions_League
28: 574076 S.L._Benfica
29: 23033 Portugal
30: 8913012 Wembley_Stadium
31: 16415 Juventus_F.C.
32: 44220 UEFA_Champions_League
33: 2273 AFC_Ajax
34: 19961 Manchester_United_F.C.
35: 9316 England
36: 18119 Liverpool_F.C.
37: 61588 King_Baudouin_Stadium
38: 3708 Brussels

1: 19961 Manchester_United_F.C.
[pem_pos=0, final_scores=0.5233918,
final_local_scores=0.47625482, pem=0.0, similarity=18.302107]
2: 16415 Juventus_F.C.
[pem_pos=0, final_scores=0.48008883,
final_local_scores=0.44032735, pem=0.0, similarity=17.22996]
3: 9239 Europe
[pem_pos=0, final_scores=0.57205707,
final_local_scores=0.6489731, pem=-0.086338006,

```
similarity=23.925884]
4: 12521 Geneva
    [pem_pos=0, final_scores=0.4189425,
final_local_scores=0.46353352, pem=-0.24772356,
similarity=19.269636]
5: 44220 UEFA_Champions_League
    [pem_pos=14, final_scores=0.21626176,
final_local_scores=0.15054, pem=-4.3849063,
similarity=32.42793]
6: 16415 Juventus_F.C.
    [pem_pos=0, final_scores=0.50756127,
final_local_scores=0.4703368, pem=0.0, similarity=18.125502]
7: 9316 England
    [pem_pos=1, final_scores=0.5276386,
final_local_scores=0.5742317, pem=-1.3527626,
similarity=28.582466]
8: 19961 Manchester_United_F.C.
    [pem_pos=0, final_scores=0.58580846,
final_local_scores=0.5444356, pem=0.0, similarity=20.336758]
9: 16415 Juventus_F.C.
    [pem_pos=0, final_scores=0.483954,
final_local_scores=0.44454944, pem=0.0, similarity=17.355955]
10: 21148 Netherlands
    [pem_pos=0, final_scores=0.2630375,
final_local_scores=0.27404904, pem=-0.7943452,
similarity=16.587656]
11: 2273 AFC_Ajax
    [pem_pos=0, final_scores=0.6532797,
final_local_scores=0.63621956, pem=0.0, similarity=23.075773]
12: 43223 Alex_Ferguson
    [pem_pos=0, final_scores=0.53559047,
final_local_scores=0.5420722, pem=0.0, similarity=20.266228]
13: 656932 SK_Rapid_Wien
    [pem_pos=0, final_scores=0.50936943,
final_local_scores=0.51525635, pem=0.0, similarity=19.46599]
14: 979542 Fenerbahçe_S.K.
    [pem_pos=0, final_scores=0.29030567,
final_local_scores=0.30270454, pem=0.0, similarity=13.12302]
15: 16415 Juventus_F.C.
    [pem_pos=0, final_scores=0.49775463,
final_local_scores=0.45962453, pem=0.0, similarity=17.805826]
16: 19961 Manchester_United_F.C.
    [pem_pos=5, final_scores=0.540384,
final_local_scores=0.49664265, pem=-3.3400002,
similarity=37.07396]
17: 19450529 Turin
    [pem_pos=0, final_scores=0.4619419,
final_local_scores=0.50277114, pem=-0.048797484,
similarity=19.358774]
18: 660592 Old_Trafford
    [pem_pos=0, final_scores=0.47259575,
final_local_scores=0.44525295, pem=0.0, similarity=17.37695]
```

19: 19961 Manchester_United_F.C.
[pem_pos=5, final_scores=0.4789456,
final_local_scores=0.42770416, pem=-3.3400002,
similarity=35.016697]
20: 9316 England
[pem_pos=1, final_scores=0.5024337,
final_local_scores=0.5466991, pem=-1.3527626,
similarity=27.76084]
21: 9239 Europe
[pem_pos=0, final_scores=0.5755722,
final_local_scores=0.6528128, pem=-0.086338006,
similarity=24.040468]
22: 44220 UEFA_Champions_League
[pem_pos=0, final_scores=0.4599318,
final_local_scores=0.41671246, pem=-0.5339415,
similarity=19.428902]
23: 23608452 Galatasaray_S.K._(football)
[pem_pos=1, final_scores=0.34737474,
final_local_scores=0.33991528, pem=-1.8990651,
similarity=24.560875]
24: 26667 Spain
[pem_pos=0, final_scores=0.47986382,
final_local_scores=0.5208248, pem=-0.18857051,
similarity=20.657639]
25: 68187 FC_Barcelona
[pem_pos=1, final_scores=0.46662855,
final_local_scores=0.4233175, pem=-1.7089969,
similarity=26.01614]
26: 68187 FC_Barcelona
[pem_pos=1, final_scores=0.4611414,
final_local_scores=0.41732365, pem=-1.7089969,
similarity=25.837273]
27: 44220 UEFA_Champions_League
[pem_pos=0, final_scores=0.42306048,
final_local_scores=0.37643617, pem=-0.5339415,
similarity=18.226974]
28: 574076 S.L._Benfica
[pem_pos=0, final_scores=0.42199224,
final_local_scores=0.39184833, pem=0.0, similarity=15.783248]
29: 23033 Portugal
[pem_pos=0, final_scores=0.527577,
final_local_scores=0.56611323, pem=-0.30942732,
similarity=22.666374]
30: 8913012 Wembley_Stadium
[pem_pos=2, final_scores=0.10639267,
final_local_scores=0.07135701, pem=-1.2725266,
similarity=13.139346]
31: 16415 Juventus_F.C.
[pem_pos=0, final_scores=0.49530047,
final_local_scores=0.4569438, pem=0.0, similarity=17.725828]
32: 44220 UEFA_Champions_League
[pem_pos=0, final_scores=0.42815095,

```
final_local_scores=0.3819967, pem=-0.5339415,
similarity=18.392914]
33: 2273 AFC_Ajax
    [pem_pos=1, final_scores=0.47477263,
final_local_scores=0.4412276, pem=-1.4356393,
similarity=25.064053]
34: 19961 Manchester_United_F.C.
    [pem_pos=5, final_scores=0.40761733,
final_local_scores=0.34978884, pem=-3.3400002,
similarity=32.691547]
35: 9316 England
    [pem_pos=1, final_scores=0.45811343,
final_local_scores=0.49828595, pem=-1.3527626,
similarity=26.316093]
36: 18119 Liverpool_F.C.
    [pem_pos=1, final_scores=0.5255922,
final_local_scores=0.4795729, pem=-1.1702383,
similarity=24.765062]
37: 61588 King_Baudouin_Stadium
    [pem_pos=1, final_scores=0.26354736,
final_local_scores=0.28044456, pem=-1.3288279,
similarity=19.68511]
38: 3708 Brussels
    [pem_pos=0, final_scores=0.6271743,
final_local_scores=0.6808245, pem=-0.025440974,
similarity=24.545227]
```

Sample: b'181_Nigeria'

```
[1 1 Nigeria 1] police kill six robbery suspects . [2 2 LAGOS
2] 1996-08-23 [3 3 Nigerian 3] police shot dead six robbery
suspects as they tried to escape from custody in the northern
city of [4 4 Sokoto 4] , the national news agency reported on
Friday . The News Agency of Nigeria ( NAN ) quoted police
spokesman Umar Shelling as saying the six were killed on
Wednesday . They had been arrested last week for stealing
800,000 naira ( $ 10,000 ) from a sheep merchant .
```

```
1: 21383 Nigeria
2: 85232 Lagos
3: 21383 Nigeria
4: 537175 Sokoto
1: 21383 Nigeria
    [pem_pos=0, final_scores=0.61805034,
final_local_scores=0.5829332, pem=-0.31455088,
similarity=23.196178]
2: 85232 Lagos
    [pem_pos=0, final_scores=0.47568113,
final_local_scores=0.45929086, pem=-0.20534219,
similarity=18.912552]
3: 21383 Nigeria
    [pem_pos=0, final_scores=0.5574444,
final_local_scores=0.5167304, pem=-0.075027294,
similarity=19.91799]
```

```
4: 537175 Sokoto
    [pem_pos=0, final_scores=0.23684193,
final_local_scores=0.2474744, pem=-0.30492392,
similarity=13.133066]

Sample: b'182_Rwandan'
[1 1 Rwandan 1] group says expulsion could be imminent . [2 2
NAIROBI 2] 1996-08-23 Repatriation of 1.1 million Rwandan
Hutu refugees announced by [3 3 Zaire 3] and [4 4 Rwanda 4]
on Thursday could start within the next few days , an exiled
Rwandan Hutu lobby group said on Friday . Innocent Butare ,
executive secretary of the Rally for the Return of Refugees
and Democracy in Rwanda ( [5 5 RDR 5] ) which says it has the
support of [6 6 Rwanda 6] 's exiled Hutus , appealed to the
international community to deter the two countries from going
ahead with what it termed a " forced and inhuman action " .
1: 25645 Rwanda
2: 21482 Nairobi
3: 76762 Democratic_Republic_of_the_Congo
4: 25645 Rwanda
5: 1405778 Republican_Rally_for_Democracy_in_Rwanda
6: 25645 Rwanda
1: 25645 Rwanda
    [pem_pos=0, final_scores=0.6235314,
final_local_scores=0.56811774, pem=0.0, similarity=21.04348]
2: 21482 Nairobi
    [pem_pos=0, final_scores=0.44606417,
final_local_scores=0.45630896, pem=0.0, similarity=17.706882]
3: 76762 Democratic_Republic_of_the_Congo
    [pem_pos=3, final_scores=0.36703247,
final_local_scores=0.33519837, pem=-2.0037596,
similarity=24.989458]
4: 25645 Rwanda
    [pem_pos=0, final_scores=0.5898811,
final_local_scores=0.5313599, pem=-0.42949,
similarity=22.282187]
5: 1405778 Republican_Rally_for_Democracy_in_Rwanda
    [pem_pos=1, final_scores=0.46266776,
final_local_scores=0.57349443, pem=-1.4234477,
similarity=28.94486]
6: 25645 Rwanda
    [pem_pos=0, final_scores=0.63669854,
final_local_scores=0.58250093, pem=-0.42949,
similarity=23.808336]

Sample: b'183_Orthodox'
Orthodox church blown up in southern [1 1 Croatia 1] . [2 2 Z
AGREB 2] 1996-08-23 Saboteurs blew up a [3 3 Serb 3] orthodox
church in southern [4 4 Croatia 4] on Friday with a blast
which also damaged four nearby homes , the state news agency
Hina reported . 1 HINA said the church in the small village
of Karin Gornji , 30 km ( 19 miles ) north of [5 5 Zadar 5] ,
```

was destroyed by the morning attack . It did not report any casualties . [6 6 Zadar 6] police said in a statement they had launched an investigation and were doing their best to find the perpetrators . 2 HINA said it was the first time an orthodox church had been blown up in the [7 7 Zadar 7] hinterland , where a large number of [8 8 Serbs 8] lived before the 1991 war over [9 9 Croatia 9] 's independence from the [10 10 Yugoslav 10] federation . The area was part of the self-styled state of [11 11 Krajina 11] proclaimed by minority [12 12 Serbs 12] in 1991 and recaptured by the [13 13 Croatian 13] army last year . Up to 200,000 [14 14 Serbs 14] fled to Bosnia and [15 15 Yugoslavia 15] , leaving [16 16 Kra jina 16] vacant and depopulated .

1: 5573 Croatia
2: 55906 Zagreb
3: 29265 Serbia
4: 5573 Croatia
5: 292128 Zadar
6: 292128 Zadar
7: 292128 Zadar
8: 59512 Serbs
9: 5573 Croatia
10: 34244 Yugoslavia
11: 526399 Republic_of_Serbian_Krajina
12: 59512 Serbs
13: 5573 Croatia
14: 59512 Serbs
15: 34244 Yugoslavia
16: 526399 Republic_of_Serbian_Krajina
1: 10601955 HINA
[winner: 16351189 Hina_Rabbani_Khar, pem_pos=6,
final_scores=-0.300889, final_local_scores=-0.20649302,
pem=-2.0794415, similarity=9.235876]
RECALL MISS: 10601955 HINA]
2: 10601955 HINA
[winner: 16351189 Hina_Rabbani_Khar, pem_pos=6,
final_scores=-0.2771777, final_local_scores=-0.18059203,
pem=-2.0794415, similarity=10.008813]
RECALL MISS: 10601955 HINA]
1: 5573 Croatia
[pem_pos=0, final_scores=0.6114411,
final_local_scores=0.5670603, pem=-0.35021794,
similarity=22.916462]
2: 55906 Zagreb
[pem_pos=0, final_scores=0.5081808,
final_local_scores=0.47070312, pem=-0.06669851,
similarity=18.49915]
3: 29265 Serbia
[pem_pos=1, final_scores=0.54892486,
final_local_scores=0.5131991, pem=-2.1726613,
similarity=31.219866]
4: 5573 Croatia

```
[pem_pos=0, final_scores=0.61490566,
final_local_scores=0.57084477, pem=-0.35021794,
similarity=23.029398]
5: 292128 Zadar
[pem_pos=0, final_scores=0.4335782,
final_local_scores=0.39670473, pem=-0.050869767,
similarity=16.204811]
6: 292128 Zadar
[pem_pos=0, final_scores=0.38351852,
final_local_scores=0.34202218, pem=-0.050869767,
similarity=14.572974]
7: 292128 Zadar
[pem_pos=0, final_scores=0.453839,
final_local_scores=0.4188366, pem=-0.050869767,
similarity=16.86527]
8: 59512 Serbs
[pem_pos=0, final_scores=0.39869803,
final_local_scores=0.34610096, pem=-0.02793329,
similarity=14.569962]
9: 5573 Croatia
[pem_pos=0, final_scores=0.6118118,
final_local_scores=0.56746525, pem=-0.35021794,
similarity=22.928547]
10: 34244 Yugoslavia
[pem_pos=0, final_scores=0.38345551,
final_local_scores=0.32821876, pem=-0.7596219,
similarity=18.015358]
11: 526399 Republic_of_Serbian_Krajina
[pem_pos=1, final_scores=0.3669297,
final_local_scores=0.3863334, pem=-0.85154414,
similarity=20.2495]
12: 59512 Serbs
[pem_pos=0, final_scores=0.39269775,
final_local_scores=0.33954653, pem=-0.02793329,
similarity=14.374364]
13: 5573 Croatia
[pem_pos=1, final_scores=0.54014015,
final_local_scores=0.48917484, pem=-1.1038951,
similarity=24.690819]
14: 59512 Serbs
[pem_pos=0, final_scores=0.3967437,
final_local_scores=0.34396616, pem=-0.02793329,
similarity=14.506254]
15: 34244 Yugoslavia
[pem_pos=0, final_scores=0.37300915,
final_local_scores=0.31680766, pem=-0.6536551,
similarity=17.098564]
16: 526399 Republic_of_Serbian_Krajina
[pem_pos=1, final_scores=0.40288073,
final_local_scores=0.4256044, pem=-0.85154414,
similarity=21.421427]
```

Sample: b'184_Hungary'

[1 1 Hungary 1] 's gross foreign debt rises in June . [2 2 BU
DAPEST 2] 1996-08-23 [3 3 Hungary 3] 's gross foreign debt
rose to \$ 27.53 billion in June from \$ 27.25 billion in May ,
the [4 4 National Bank of Hungary 4] (NBH) said on Friday .
FIGURES IN \$ MILLION June 1996 May 1996 Gross foreign debt
27,535.5 27,246.5 International reserves and other foreign
assets 13,256.5 12,855.7 Net foreign debt 14,278.9 14,390.7
Net foreign debt of the government and NBH 9,510.9 10,056.4
-- [5 5 Budapest 5] newsroom (36 1) 266 2410

1: 13275 Hungary
2: 36787 Budapest
3: 13275 Hungary
4: 2395098 Hungarian_National_Bank
5: 36787 Budapest
1: 13275 Hungary
[pem_pos=0, final_scores=0.6169298,
final_local_scores=0.57978237, pem=-0.32577527,
similarity=23.163193]
2: 36787 Budapest
[pem_pos=0, final_scores=0.54360455,
final_local_scores=0.49482232, pem=-0.030229695,
similarity=19.020592]
3: 13275 Hungary
[pem_pos=0, final_scores=0.6413271,
final_local_scores=0.60643256, pem=-0.32577527,
similarity=23.958487]
4: 2395098 Hungarian_National_Bank
[pem_pos=0, final_scores=0.45223385,
final_local_scores=0.5554909, pem=0.0, similarity=20.66667]
5: 36787 Budapest
[pem_pos=0, final_scores=0.6390173,
final_local_scores=0.59904623, pem=-0.030229695,
similarity=22.13084]

Sample: b'185_Germany'

[1 1 Germany 1] , [2 2 Poland 2] tighten cooperation against
crime . [3 3 WARSAW 3] 1996-08-23 [4 4 Germany 4] and [5 5 Po
land 5] agreed on Friday to tighten cooperation between their
intelligence services in fighting international organised
crime , [6 6 PAP 6] news agency reported . [7 7 Interior 7]
Minister [8 8 Zbigniew Siemiatkowski 8] and Bernd
Schmidbauer , [9 9 German 9] intelligence co-ordinator in [10
10 Helmut Kohl 10] 's chancellery , sealed the closer links
during talks in [11 11 Warsaw 11] . Ministry spokesman
Ryszard Hincza told the [12 12 Polish 12] agency the services
would work together against mafia-style groups , drug
smuggling and illegal trade in arms and radioactive
materials .

1: 11867 Germany
2: 22936 Poland
3: 32908 Warsaw

4: 11867 Germany
5: 22936 Poland
6: 1853823 Polish_Press_Agency
7: 354620 Interior_ministry
8: 8356016 Zbigniew_Siemiątkowski
9: 11867 Germany
10: 41896 Helmut_Kohl
11: 32908 Warsaw
12: 22936 Poland
1: 11867 Germany
[pem_pos=0, final_scores=0.7560745,
final_local_scores=0.73680544, pem=-0.15529114,
similarity=26.921955]
2: 22936 Poland
[pem_pos=0, final_scores=0.60168624,
final_local_scores=0.55626315, pem=-0.15333635,
similarity=21.52358]
3: 32908 Warsaw
[pem_pos=0, final_scores=0.4960035,
final_local_scores=0.46201062, pem=-0.14300999,
similarity=18.654741]
4: 11867 Germany
[pem_pos=0, final_scores=0.838602,
final_local_scores=0.82695425, pem=-0.15529114,
similarity=29.612175]
5: 22936 Poland
[pem_pos=0, final_scores=0.5933359,
final_local_scores=0.5471417, pem=-0.15333635,
similarity=21.251379]
6: 1853823 Polish_Press_Agency
[pem_pos=4, final_scores=0.5298267,
final_local_scores=0.62934214, pem=-2.656357,
similarity=37.316223]
7: 354620 Interior_ministry
[pem_pos=3, final_scores=0.19075982,
final_local_scores=0.22940296, pem=-2.5722044,
similarity=24.923601]
8: 8356016 Zbigniew_Siemiątkowski
[pem_pos=0, final_scores=0.57173836,
final_local_scores=0.71709913, pem=0.0, similarity=25.489382]
9: 11867 Germany
[pem_pos=1, final_scores=0.8347841,
final_local_scores=0.82278377, pem=-1.0465823,
similarity=34.3347]
10: 41896 Helmut_Kohl
[pem_pos=0, final_scores=0.73581284,
final_local_scores=0.80634284, pem=0.0, similarity=28.152592]
11: 32908 Warsaw
[pem_pos=0, final_scores=0.53743136,
final_local_scores=0.50726426, pem=-0.14300999,
similarity=20.005201]
12: 22936 Poland

[pem_pos=0, final_scores=0.6008768,
final_local_scores=0.555379, pem=-0.81518245,
similarity=25.096418]

Sample: b'186_Russians'
[1 1 Russians 1] , [2 2 Chechens 2] say observing [3 3 Grozny
3] ceasefire . [4 4 GROZNY 4] , [5 5 Russia 5] 1996-08-23
Rebel fighters and [6 6 Russian 6] soldiers said a ceasefire
effective at noon (0800 [7 7 GMT 7]) on Friday was being
generally observed , although scattered gunfire echoed
through the [8 8 Chechen 8] capital [9 9 Grozny 9] . The [10 1
0 Russian 10] army said earlier it was preparing to withdraw
from the rebel-dominated southern mountains of the region as
part of the peace deal reached with separatists on Thursday .
" There has been some shooting from their side but it has
been relatively quiet , " said fighter Aslan Shabazov , a
bearded man wearing a white t-shirt and camouflage trousers .
Soon after he spoke another burst of gunfire rocked the
courtyard where the rebels had set up their base and a
captured [11 11 Russian 11] [12 12 T-72 12] tank roared out
to investigate . The separatists , who swept into [13 13 Groz
ny 13] on August 6 , still control large areas of the centre
of town , and [14 14 Russian 14] soldiers are based at
checkpoints on the approach roads . " The ceasefire is being
observed , " said woman soldier Svetlana Goncharova , 35 ,
short dark hair poking out from under a peaked camouflage
cap . A few helicopters flew overhead , firing off flares ,
but there was no shooting from the air . The truce , the
latest of several , was agreed in talks on Thursday between [1
5 15 Russian 15] peacemaker [16 16 Alexander Lebed 16] and
rebel chief-of-staff [17 17 Aslan Maskhadov 17] . The two
also agreed to set up joint patrols in [18 18 Grozny 18] ,
but Goncharova said she was sceptical about whether this
could work . " We have to try it , but I doubt if this is
possible with the separatists , " she said .

1: 25391 Russia
2: 6095 Chechnya
3: 140790 Grozny
4: 140790 Grozny
5: 25391 Russia
6: 25391 Russia
7: 12701 Greenwich_Mean_Time
8: 6095 Chechnya
9: 140790 Grozny
10: 25391 Russia
11: 25391 Russia
12: 186191 T-72
13: 140790 Grozny
14: 25391 Russia
15: 25391 Russia
16: 431579 Alexander_Lebed
17: 340963 Aslan_Maskhadov

```
18: 140790 Grozny
1: 25391 Russia
    [pem_pos=1, final_scores=0.66296655,
final_local_scores=0.6404895, pem=-1.8802736,
similarity=33.428417]
2: 6095 Chechnya
    [pem_pos=2, final_scores=0.4489097,
final_local_scores=0.4179899, pem=-0.96860826,
similarity=21.830807]
3: 140790 Grozny
    [pem_pos=0, final_scores=0.5516234,
final_local_scores=0.5230499, pem=0.0, similarity=19.698565]
4: 140790 Grozny
    [pem_pos=0, final_scores=0.55084085,
final_local_scores=0.5221951, pem=0.0, similarity=19.673056]
5: 25391 Russia
    [pem_pos=0, final_scores=0.85381085,
final_local_scores=0.84895796, pem=-0.20299491,
similarity=30.52823]
6: 25391 Russia
    [pem_pos=1, final_scores=0.9218449,
final_local_scores=0.9232748, pem=-1.0515022,
similarity=37.36031]
7: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.64191335,
final_local_scores=0.76161164, pem=0.0, similarity=26.817724]
8: 6095 Chechnya
    [pem_pos=2, final_scores=0.49325854,
final_local_scores=0.46643424, pem=-0.8617233,
similarity=22.695227]
9: 140790 Grozny
    [pem_pos=0, final_scores=0.50266117,
final_local_scores=0.4695661, pem=0.0, similarity=18.102503]
10: 25391 Russia
    [pem_pos=1, final_scores=0.8484817,
final_local_scores=0.84313667, pem=-1.0515022,
similarity=34.968826]
11: 25391 Russia
    [pem_pos=1, final_scores=0.75194293,
final_local_scores=0.7376827, pem=-1.0515022,
similarity=31.821867]
12: 186191 T-72
    [pem_pos=0, final_scores=0.33164185,
final_local_scores=0.4192211, pem=0.0, similarity=16.600107]
13: 140790 Grozny
    [pem_pos=0, final_scores=0.59659964,
final_local_scores=0.5721796, pem=0.0, similarity=21.164694]
14: 25391 Russia
    [pem_pos=1, final_scores=0.896215,
final_local_scores=0.89527804, pem=-1.0515022,
similarity=36.52483]
15: 25391 Russia
```

```
[pem_pos=1, final_scores=0.822213,
final_local_scores=0.81444204, pem=-1.0515022,
similarity=34.11252]
16: 431579 Alexander_Lebed
    [pem_pos=0, final_scores=0.8351986,
final_local_scores=0.9149479, pem=0.0, similarity=31.393585]
17: 340963 Aslan_Maskhadov
    [pem_pos=0, final_scores=0.55177695,
final_local_scores=0.5947379, pem=0.0, similarity=21.837877]
18: 140790 Grozny
    [pem_pos=0, final_scores=0.5537227,
final_local_scores=0.52534306, pem=0.0, similarity=19.766998]
```

Sample: b'187_WEATHER'

WEATHER - Conditions at CIS airports - August 23 . [1 1 MOSCOW 1] 1996-08-23 No closures of airports in the Commonwealth of Independent States are expected on August 24 and August 25 , the Russian Weather Service said on Friday . -- Moscow Newsroom +7095 941 8520

1: 19004 Moscow

1: 19004 Moscow

```
[pem_pos=0, final_scores=0.37787992,
final_local_scores=0.6853636, pem=-0.106582455,
similarity=25.121943]
```

Sample: b'188_Granic'

Granic arrives to sign Croatia-Yugoslavia treaty . [1 1 BELGRADE 1] 1996-08-23 [2 2 Yugoslavia 2] and [3 3 Croatia 3] were poised on Friday to sign a landmark normalisation treaty ending five years of tensions and paving way for stabilisation in the [4 4 Balkans 4] . [5 5 Croatian 5] Foreign Minister [6 6 Mate Granic 6] landed at [7 7 Belgrade 7] airport aboard a [8 8 Croatian 8] government jet on Friday morning for talks with his [9 9 Yugoslav 9] counterparts and a signing ceremony expected around noon (1000 [10 10 GMT 10]) . On Thursday the [11 11 Yugoslav 11] government endorsed the text of the agreement on normalising relations between the two countries , the [12 12 Yugoslav 12] news agency [13 13 Tanjug 13] said . " The government assessed the agreement as a crucial step to resolving the [14 14 Yugoslav 14] crisis , ensuring the restoration of peace in former [15 15 Yugoslavia 15] , " it said . Last-minute talks this week on the legal fine print finally cleared the way for a treaty based on mutual recognition within internationally recognised borders and the establishment of diplomatic relations , diplomats said . The pact ends five years of hostility after [16 16 Croatia 16] 's secession from federal [17 17 Yugoslavia 17] . [18 18 Western 18] powers regard diplomatic normalisation between [19 19 Croatia 19] and [20 20 Serbia 20] , twin pillars of the old multinational federal [21 21 Yugoslavia 21] , as a crucial step towards a lasting peace in the [22 22 Balkans 22]

```
1: 55904 Belgrade
2: 34244 Yugoslavia
3: 5573 Croatia
4: 4829 Balkans
5: 5573 Croatia
6: 1192234 Mate_Granić
7: 55904 Belgrade
8: 5573 Croatia
9: 34244 Yugoslavia
10: 12701 Greenwich_Mean_Time
11: 34244 Yugoslavia
12: 34244 Yugoslavia
13: 4151905 Tanjug
14: 34244 Yugoslavia
15: 34244 Yugoslavia
16: 5573 Croatia
17: 34244 Yugoslavia
18: 21208200 Western_world
19: 5573 Croatia
20: 29265 Serbia
21: 34244 Yugoslavia
22: 4829 Balkans
1: 55904 Belgrade
    [pem_pos=0, final_scores=0.4947614,
final_local_scores=0.45306635, pem=-0.05193398,
similarity=17.892542]
2: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.4339077,
final_local_scores=0.35763207, pem=-0.6536551,
similarity=18.316845]
3: 5573 Croatia
    [pem_pos=0, final_scores=0.5565161,
final_local_scores=0.5107551, pem=-0.35021794,
similarity=21.236204]
4: 4829 Balkans
    [pem_pos=0, final_scores=0.46833986,
final_local_scores=0.46800768, pem=0.0, similarity=18.055996]
5: 5573 Croatia
    [pem_pos=1, final_scores=0.547772,
final_local_scores=0.5012035, pem=-1.1038951,
similarity=25.049778]
6: 1192234 Mate_Granić
    [pem_pos=0, final_scores=0.5319688,
final_local_scores=0.65311813, pem=0.0, similarity=23.580061]
7: 55904 Belgrade
    [pem_pos=0, final_scores=0.45320517,
final_local_scores=0.40767252, pem=-0.05193398,
similarity=16.537899]
8: 5573 Croatia
    [pem_pos=1, final_scores=0.54681265,
final_local_scores=0.50015557, pem=-1.1038951,
similarity=25.018505]
```

```
9: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.3840391,
final_local_scores=0.30315816, pem=-0.7596219,
similarity=17.2675]
10: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.63421386,
final_local_scores=0.76115835, pem=0.0, similarity=26.804195]
11: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.3943104,
final_local_scores=0.31437802, pem=-0.7596219,
similarity=17.602322]
12: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.41373092,
final_local_scores=0.33559203, pem=-0.7596219,
similarity=18.23539]
13: 4151905 Tanjug
    [pem_pos=0, final_scores=0.49806565,
final_local_scores=0.589874, pem=0.0, similarity=21.69273]
14: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.39629382,
final_local_scores=0.3165446, pem=-0.7596219,
similarity=17.666977]
15: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.46533364,
final_local_scores=0.39196014, pem=-0.6536551,
similarity=19.341265]
16: 5573 Croatia
    [pem_pos=0, final_scores=0.6164889,
final_local_scores=0.5762663, pem=-0.35021794,
similarity=23.191187]
17: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.52793485,
final_local_scores=0.46034247, pem=-0.6536551,
similarity=21.38193]
18: 21208200 Western_world
    [pem_pos=0, final_scores=0.30265707,
final_local_scores=0.3396098, pem=-1.5917408,
similarity=22.88048]
19: 5573 Croatia
    [pem_pos=0, final_scores=0.6204731,
final_local_scores=0.5806184, pem=-0.35021794,
similarity=23.321062]
20: 29265 Serbia
    [pem_pos=0, final_scores=0.63011986,
final_local_scores=0.60473144, pem=-0.34077588,
similarity=23.989298]
21: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.5083968,
final_local_scores=0.43900007, pem=-0.6536551,
similarity=20.74503]
22: 4829 Balkans
    [pem_pos=0, final_scores=0.5261679,
```

```
final_local_scores=0.5311761, pem=0.0, similarity=19.941067]
```

Sample: b'189_Ecuador'

[1 1 Ecuador 1] president to lunch with ethnic [2 2 Indians 2] . [3 3 QUITO 3] , [4 4 Ecuador 4] 1996-08-23 [5 5 Ecuador 5] 's President [6 6 Abdala Bucaram 6] has announced he will hold regular lunches in his presidential palace for members of the country 's different ethnic groups as of next week . " It was about time for the [7 7 Indians 7] , the blacks and the mixed-bloods to begin eating in the palace with their president because this is not a palace exclusively for the potentates and ambassadors and protocol , " [8 8 Bucaram 8] said late on Thursday . " In these weekly lunches we are going to get to know the problems of the [9 9 Indian 9] , mixed-race , black and peasant sectors , " he said . He has invited 35 [10 10 Indian 10] leaders to lunch next Tuesday . [11 11 Bucaram 11] , who was elected on a populist platform last month , also plans to create a ministry for ethnic cultures . The Andean nation 's population of 11.4 million is 47 percent indigenous .

1: 9334 Ecuador

2: 2402376 Indigenous_peoples_of_the_Americas

3: 209982 Quito

4: 9334 Ecuador

5: 9334 Ecuador

6: 1750198 Abdalá_Bucaram

7: 2402376 Indigenous_peoples_of_the_Americas

8: 1750198 Abdalá_Bucaram

9: 2402376 Indigenous_peoples_of_the_Americas

10: 2402376 Indigenous_peoples_of_the_Americas

11: 1750198 Abdalá_Bucaram

1: 9334 Ecuador

[pem_pos=0, final_scores=0.6620092,
final_local_scores=0.71052486, pem=-0.21404564,
similarity=26.457205]

2: 2402376 Indigenous_peoples_of_the_Americas

[pem_pos=1, final_scores=0.3642767,
final_local_scores=0.3139253, pem=-1.6504525,
similarity=22.433287]

3: 209982 Quito

[pem_pos=0, final_scores=0.4611389,
final_local_scores=0.47941196, pem=0.0, similarity=18.396322]

4: 9334 Ecuador

[pem_pos=0, final_scores=0.6140334,
final_local_scores=0.6581186, pem=-0.21404564,
similarity=24.893297]

5: 9334 Ecuador

[pem_pos=0, final_scores=0.5685699,
final_local_scores=0.6084566, pem=-0.21404564,
similarity=23.411285]

6: 1750198 Abdalá_Bucaram

[pem_pos=0, final_scores=0.59630376,

```
final_local_scores=0.705096, pem=0.0, similarity=25.131184]
7: 2402376 Indigenous_peoples_of_the_Americas
    [pem_pos=1, final_scores=0.316081,
final_local_scores=0.26127878, pem=-1.6504525,
similarity=20.86221]
8: 1750198 Abdalá_Bucaram
    [pem_pos=0, final_scores=0.35335964,
final_local_scores=0.4397164, pem=0.0, similarity=17.211727]
9: 2402376 Indigenous_peoples_of_the_Americas
    [pem_pos=5, final_scores=0.25285584,
final_local_scores=0.1922149, pem=-3.522857,
similarity=28.983633]
10: 2402376 Indigenous_peoples_of_the_Americas
    [pem_pos=5, final_scores=0.23831157,
final_local_scores=0.17632753, pem=-3.522857,
similarity=28.509521]
11: 1750198 Abdalá_Bucaram
    [pem_pos=0, final_scores=0.5256766,
final_local_scores=0.62794656, pem=0.0, similarity=22.82889]
```

Sample: b'190_Brazil'

[1 1 Brazil 1] to use hovercrafts for [2 2 Amazon 2]
travel . [3 3 BRASILIA 3] 1996-08-22 Hovercrafts will soon be
plying the waters of the [4 4 Amazon 4] in a bid to reduce
the difficulties of transportation on the vast [5 5 Brazilian
5] waterway , the government said on Thursday . Two Russian-
built hovercrafts , capable of carrying up to 50 tons each ,
will begin ferrying passengers and cargo up and down the huge
river from its mouth at [6 6 Belem 6] by the end of the
year , [7 7 Brazil 7] 's Amazon Affairs Department said in a
statement . The use of riverways in the region has been made
a priority under a government plan for the [8 8 Amazon 8] and
the high-speed hovercraft will help reduce the time involved
in travelling often massive distances , it said .

```
1: 3383 Brazil
2: 1701 Amazon_River
3: 4752 Brasília
4: 1701 Amazon_River
5: 3383 Brazil
6: 409573 Belém
7: 3383 Brazil
8: 1701 Amazon_River
1: 3383 Brazil
    [pem_pos=0, final_scores=0.65763444,
final_local_scores=0.62950414, pem=-0.21313342,
similarity=24.034424]
2: 1701 Amazon_River
    [pem_pos=1, final_scores=0.39406854,
final_local_scores=0.37000525, pem=-1.9012599,
similarity=25.470755]
3: 4752 Brasília
    [pem_pos=0, final_scores=0.45317084,
```

```
final_local_scores=0.46695775, pem=0.0, similarity=18.024664]
4: 1701 Amazon_River
    [pem_pos=1, final_scores=0.48617762,
final_local_scores=0.4706205, pem=-1.9012599,
similarity=28.473314]
5: 3383 Brazil
    [pem_pos=0, final_scores=0.5496769,
final_local_scores=0.51157695, pem=-0.2733083,
similarity=20.842482]
6: 409573 Belém
    [pem_pos=0, final_scores=0.28834277,
final_local_scores=0.27088875, pem=-0.2292848,
similarity=13.42046]
7: 3383 Brazil
    [pem_pos=0, final_scores=0.74023247,
final_local_scores=0.71972996, pem=-0.21313342,
similarity=26.726942]
8: 1701 Amazon_River
    [pem_pos=1, final_scores=0.4396714,
final_local_scores=0.41981947, pem=-1.9012599,
similarity=26.95731]
```

Sample: b'191_HK'

HK 's [1 1 Tsang 1] to visit [2 2 Indonesia 2] , [3 3 New Zealand 3] . [4 4 HONG KONG 4] 1996-08-23 [5 5 Hong Kong 5] Financial Secretary [6 6 Donald Tsang 6] will visit [7 7 Indonesia 7] and [8 8 New Zealand 8] from August 25 to 31 , the government said on Friday . In [9 9 Jakarta 9] , [10 10 Tsang 10] will meet President [11 11 Suharto 11] , Minister of [12 12 Finance 12] Mar'ie Muhammad , Minister of Foreign Affairs [13 13 Ali Alatas 13] and Minister of Trade and Industry Tungky Ariwibowo . On his [14 14 New Zealand 14] leg from August 29 , [15 15 Tsang 15] will meet Prime Minister [16 16 Jim Bolger 16] , Deputy Prime Minister [17 17 Don McKinnon 17] and Minister of [18 18 Finance 18] [19 19 Bill Birch 19] .

```
1: 336773 Donald_Tsang
2: 14579 Indonesia
3: 4913064 New_Zealand
4: 13404 Hong_Kong
5: 13404 Hong_Kong
6: 336773 Donald_Tsang
7: 14579 Indonesia
8: 4913064 New_Zealand
9: 16275 Jakarta
10: 336773 Donald_Tsang
11: 217513 Suharto
12: 354563 Finance_minister
13: 2488298 Ali_Alatas
14: 4913064 New_Zealand
15: 336773 Donald_Tsang
16: 362227 Jim_Bolger
17: 331647 Don_McKinnon
```

```
18: 354563 Finance_minister
19: 595530 Bill_Birch
1: 336773 Donald_Tsang
    [pem_pos=0, final_scores=1.055804,
final_local_scores=1.0965476, pem=0.0, similarity=36.812878]
2: 14579 Indonesia
    [pem_pos=0, final_scores=0.548842,
final_local_scores=0.5870569, pem=-0.23151328,
similarity=22.867666]
3: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.5867143,
final_local_scores=0.62553954, pem=-0.23795462,
similarity=24.051094]
4: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.6479164,
final_local_scores=0.6455662, pem=-0.22402541,
similarity=24.572983]
5: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5670635,
final_local_scores=0.5572468, pem=-0.22402541,
similarity=21.937355]
6: 336773 Donald_Tsang
    [pem_pos=0, final_scores=0.74797606,
final_local_scores=0.7524691, pem=0.0, similarity=26.544893]
7: 14579 Indonesia
    [pem_pos=0, final_scores=0.56792223,
final_local_scores=0.60789907, pem=-0.23151328,
similarity=23.48964]
8: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.5510426,
final_local_scores=0.5865736, pem=-0.23795462,
similarity=22.888273]
9: 16275 Jakarta
    [pem_pos=0, final_scores=0.42281944,
final_local_scores=0.45114225, pem=0.0, similarity=17.552698]
10: 336773 Donald_Tsang
    [pem_pos=0, final_scores=1.0588487,
final_local_scores=1.0998735, pem=0.0, similarity=36.912132]
11: 217513 Suharto
    [pem_pos=0, final_scores=0.36406654,
final_local_scores=0.37821257, pem=0.0, similarity=15.376331]
12: 354563 Finance_minister
    [pem_pos=12, final_scores=0.51691234,
final_local_scores=0.54446054, pem=-3.2094803,
similarity=37.791153]
13: 2488298 Ali_Alatas
    [pem_pos=0, final_scores=0.6184474,
final_local_scores=0.7218787, pem=0.0, similarity=25.632013]
14: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.5786013,
final_local_scores=0.6166773, pem=-0.23795462,
similarity=23.786627]
```

```
15: 336773 Donald_Tsang
    [pem_pos=0, final_scores=1.0578338,
final_local_scores=1.0987649, pem=0.0, similarity=36.879047]
16: 362227 Jim_Bolger
    [pem_pos=0, final_scores=0.5752623,
final_local_scores=0.6518964, pem=0.0, similarity=23.543602]
17: 331647 Don_McKinnon
    [pem_pos=0, final_scores=0.48848832,
final_local_scores=0.53243357, pem=0.0, similarity=19.978592]
18: 354563 Finance_minister
    [pem_pos=12, final_scores=0.42504722,
final_local_scores=0.44411182, pem=-3.2094803,
similarity=34.796547]
19: 595530 Bill_Birch
    [pem_pos=0, final_scores=0.38767695,
final_local_scores=0.459491, pem=0.0, similarity=17.801842]
```

Sample: b'192_Jordan'

[1 1 Jordan 1] expels [2 2 Iraqi 2] diplomat after bread riots . Rana Sabbagh [3 3 AMMAN 3] 1996-08-23 [4 4 Jordan 4] , which has blamed [5 5 Iraq 5] for bread riots last week , has asked an [6 6 Iraqi 6] diplomat to leave , official and diplomatic sources said on Friday . The main Friday prayers in southern [7 7 Jordan 7] that were the starting point for the riots a week ago passed peacefully under tight security imposed by the army with only brief demonstrations reported . Adel Ibrahim , the [8 8 Iraqi 8] embassy 's press attache , was asked to leave " because he was carrying out duties incompatible with diplomatic norms " , one source told [9 9 Reuters 9] , implying he was accused of spying . Ibrahim told [10 10 Reuters 10] by telephone from his embassy office in [11 11 Amman 11] that he " had not been notified " of any explosion order . The government declined official comment . Ibrahim 's assistant , Hussein Khalaf , was expelled earlier this year for similar reasons amid rising tension in bilateral ties after King [12 12 Hussein 12] began calling for change in [13 13 Baghdad 13] following top [14 14 Iraqi 14] defections in August 1995 . [15 15 Iraq 15] retaliated then by expelling a junior administrator working in the [16 16 Jordanian 16] embassy in [17 17 Baghdad 17] but has continued its policy of trying to avoid public conflicts with [18 18 Jordan 18] -- its only secure route to the rest of the world . [19 19 Jordan 19] has accused [20 20 Iraq 20] and a local pro-Baghdad party for the country 's worst unrest in seven years which erupted after it almost doubled the prices of bread last week under radical economic reforms agreed with the [21 21 International Monetary Fund 21] . In [2 22 Karak 22] , where two days of riots flared last Friday , a few hundred young men lingered outside Omari mosque on leaving , shouting slogans for about 15 minutes . " Disperse , abstain from forming groups and help maintain order , " army officers , who has enforced a loose curfew

since the riots , told the crowd through loudspeakers . The men shouted " [23 23 Allahu 23] Akbar " ([24 24 God 24] is Greatest) as a former [25 25 Islamist 25] deputy , Ahmed Kafawin , told soldiers the crowd would not cause trouble . He had earlier mounted the mosque 's pulpit to demand release of detainees , an end to raids on houses and the cancelling of the bread price rises . Armoured cars had patrolled streets in [26 26 Karak 26] , traditional bastion of communist ideology and [27 27 Baath 27] socialism that swept the region in the 1950s , and guarded entrances to the hill-top city famed for its Crusader castle before the prayers . There was also heavy security in the crowded centre of [28 28 Amman 28] , where smaller clashes had erupted last Saturday , but Friday prayers at the main mosque ended quietly as police in full riot gear looked on . The Jordanian Arab Socialist Baath Party , which has one deputy in the 80-seat lower house of parliament , has denied involvement in unrest which it blamed on government policies and rising economic hardship . Government attempts to link the rioting to foreign influence has been treated with derision by those in the streets who blame the protests on severe economic hardships .

1: 7515964 Jordan
2: 7515928 Iraq
3: 45007 Amman
4: 7515964 Jordan
5: 7515928 Iraq
6: 7515928 Iraq
7: 7515964 Jordan
8: 7515928 Iraq
9: 18998750 Reuters
10: 18998750 Reuters
11: 45007 Amman
12: 80608 Hussein_of_Jordan
13: 4492 Baghdad
14: 7515928 Iraq
15: 7515928 Iraq
16: 7515964 Jordan
17: 4492 Baghdad
18: 7515964 Jordan
19: 7515964 Jordan
20: 7515928 Iraq
21: 15251 International_Monetary_Fund
22: 307576 Al_Karak
23: 740 Allah
24: 5042765 God
25: 6037917 Islam
26: 307576 Al_Karak
27: 166957 Ba'ath_Party
28: 45007 Amman
1: 7515964 Jordan
[pem_pos=0, final_scores=0.66493773,
final_local_scores=0.6218533, pem=-0.53647494,

```
similarity=25.56449]
2: 7515928 Iraq
    [pem_pos=0, final_scores=0.5361268,
final_local_scores=0.47999173, pem=-0.48469657,
similarity=21.049479]
3: 45007 Amman
    [pem_pos=0, final_scores=0.55693364,
final_local_scores=0.5309851, pem=0.0, similarity=19.935368]
4: 7515964 Jordan
    [pem_pos=0, final_scores=0.57373327,
final_local_scores=0.5222263, pem=-0.53647494,
similarity=22.591421]
5: 7515928 Iraq
    [pem_pos=0, final_scores=0.62167543,
final_local_scores=0.5734406, pem=-0.25379938,
similarity=22.582525]
6: 7515928 Iraq
    [pem_pos=0, final_scores=0.5888733,
final_local_scores=0.5376093, pem=-0.48469657,
similarity=22.768902]
7: 7515964 Jordan
    [pem_pos=0, final_scores=0.55531615,
final_local_scores=0.5021084, pem=-0.53647494,
similarity=21.991062]
8: 7515928 Iraq
    [pem_pos=0, final_scores=0.5952852,
final_local_scores=0.5446133, pem=-0.48469657,
similarity=22.977915]
9: 18998750 Reuters
    [pem_pos=0, final_scores=0.4186585,
final_local_scores=0.44063073, pem=0.0, similarity=17.239014]
10: 18998750 Reuters
    [pem_pos=0, final_scores=0.36893278,
final_local_scores=0.3863129, pem=0.0, similarity=15.618061]
11: 45007 Amman
    [pem_pos=0, final_scores=0.37672502,
final_local_scores=0.33413455, pem=0.0, similarity=14.060954]
12: 80608 Hussein_of_Jordan
    [pem_pos=0, final_scores=0.30272943,
final_local_scores=0.2938565, pem=-1.2591617,
similarity=19.706495]
13: 4492 Baghdad
    [pem_pos=0, final_scores=0.37514162,
final_local_scores=0.3162898, pem=0.0, similarity=13.528432]
14: 7515928 Iraq
    [pem_pos=0, final_scores=0.5532207,
final_local_scores=0.4986642, pem=-0.48469657,
similarity=21.606703]
15: 7515928 Iraq
    [pem_pos=0, final_scores=0.5388922,
final_local_scores=0.4830125, pem=-0.25379938,
similarity=19.88397]
```

```
16: 7515964 Jordan
    [pem_pos=0, final_scores=0.5810966,
final_local_scores=0.5302696, pem=-0.18608934,
similarity=20.925999]
17: 4492 Baghdad
    [pem_pos=0, final_scores=0.3154533,
final_local_scores=0.2510894, pem=0.0, similarity=11.582722]
18: 7515964 Jordan
    [pem_pos=0, final_scores=0.5969931,
final_local_scores=0.5476341, pem=-0.53647494,
similarity=23.349642]
19: 7515964 Jordan
    [pem_pos=0, final_scores=0.6365881,
final_local_scores=0.59088564, pem=-0.53647494,
similarity=24.640352]
20: 7515928 Iraq
    [pem_pos=0, final_scores=0.66468054,
final_local_scores=0.62041724, pem=-0.25379938,
similarity=23.984402]
21: 15251 International_Monetary_Fund
    [pem_pos=0, final_scores=0.3719321,
final_local_scores=0.45632404, pem=0.0, similarity=17.707333]
22: 307576 Al_Karak
    [pem_pos=0, final_scores=0.36206847,
final_local_scores=0.36392313, pem=-0.6855135,
similarity=18.677834]
23: 740 Allah
    [pem_pos=0, final_scores=0.5606572,
final_local_scores=0.6535391, pem=0.0, similarity=23.592625]
24: 5042765 God
    [pem_pos=0, final_scores=0.63055754,
final_local_scores=0.73876077, pem=-0.18657415,
similarity=27.150429]
25: 6037917 Islam
    [pem_pos=1, final_scores=0.30429018,
final_local_scores=0.29973504, pem=-1.3455799,
similarity=20.351877]
26: 307576 Al_Karak
    [pem_pos=0, final_scores=0.45784193,
final_local_scores=0.4685411, pem=-0.6855135,
similarity=21.799843]
27: 166957 Ba'ath_Party
    [pem_pos=0, final_scores=0.42688233,
final_local_scores=0.44525987, pem=0.0, similarity=17.377155]
28: 45007 Amman
    [pem_pos=0, final_scores=0.47186333,
final_local_scores=0.43805873, pem=0.0, similarity=17.16226]
```

Sample: b'193_Turkey'

Turkey says killed 17 [1 1 Kurd 1] rebels in clashes . [2 2 A NKARA 2] 1996-08-23 1 Turkish troops have killed 17 [3 3 Kurd ish 3] rebels in recent clashes in the southeast of the

country , the state-run [4 4 Anatolian 4] news agency said on Friday . Two security officials and two state-paid village guards were killed in the fighting with [5 5 Kurdistan Workers Party 5] ([6 6 PKK 6]) guerrillas , the agency quoted the emergency rule governor 's office as saying . Eight of the rebels were killed in [7 7 Van 7] province , five in [8 8 Sirnak 8] and four in [9 9 Hakkari 9] . The agency did not say when the clashes took place . More than 20,000 people have died in the [10 10 PKK 10] 's 12-year-old fight for independence or autonomy in southeastern Turkey . Three people , including two village guards , died when a mine planted by [11 11 PKK 11] rebels exploded on a road in the southeast , [12 12 Anatolian 12] reported earlier . It said a taxi carrying the guards , members of a mostly [1 2 Kurdish 1] militia which fights the [13 13 PKK 13] , hit the mine in the province of [2 Diyarbakir 2] .

1: 17068 Kurdish_people
2: 802 Ankara
3: 17068 Kurdish_people
4: 854 Anatolia
5: 69680 Kurdistan_Workers'_Party
6: 69680 Kurdistan_Workers'_Party
7: 63685 Van,_Turkey
8: 10541472 Şırnak_Province
9: 1372482 Hakkâri_Province
10: 69680 Kurdistan_Workers'_Party
11: 69680 Kurdistan_Workers'_Party
12: 854 Anatolia
13: 69680 Kurdistan_Workers'_Party
1: 30205 Turkish_Armed_Forces
[winner: 29992 Turkish_language, pem_pos=0,
final_scores=-0.16350067, final_local_scores=-0.20324743,
pem=-0.7274845, similarity=1.9805781]
RECALL MISS: 30205 Turkish_Armed_Forces]
2: 12545276 Kurdish_nationalism
[winner: 17068 Kurdish_people, pem_pos=1,
final_scores=0.07859047, final_local_scores=-0.0016869903,
pem=-0.713598, similarity=7.920027]
gt: 12545276 Kurdish_nationalism, pem_pos=3,
final_scores=-0.031001464, final_local_scores=-0.067744374,
pem=-2.7281096, similarity=16.903969]
1: 17068 Kurdish_people
[pem_pos=0, final_scores=0.18615727,
final_local_scores=0.12400305, pem=0.0, similarity=7.7902117]
2: 802 Ankara
[pem_pos=0, final_scores=0.38218802,
final_local_scores=0.38569176, pem=0.0, similarity=15.5995245]
3: 17068 Kurdish_people
[pem_pos=1, final_scores=0.11990665,
final_local_scores=0.05163431, pem=-0.713598,
similarity=9.511241]
4: 854 Anatolia

[pem_pos=0, final_scores=0.3970585,
final_local_scores=0.39412266, pem=-0.23601569,
similarity=17.13461]
5: 69680 Kurdistan_Workers'_Party
[pem_pos=0, final_scores=0.60326207,
final_local_scores=0.58519065, pem=0.0, similarity=21.552969]
6: 69680 Kurdistan_Workers'_Party
[pem_pos=0, final_scores=0.56530505,
final_local_scores=0.54372835, pem=0.0, similarity=20.31565]
7: 63685 Van,_Turkey
[pem_pos=3, final_scores=0.40051728,
final_local_scores=0.38536668, pem=-3.4721067,
similarity=34.47168]
8: 10541472 Şırnak_Province
[pem_pos=1, final_scores=0.3715276,
final_local_scores=0.4220075, pem=-0.68484837,
similarity=20.40757]
9: 1372482 Hakkâri_Province
[pem_pos=1, final_scores=0.14157192,
final_local_scores=0.14657778, pem=-0.47937384,
similarity=11.070795]
10: 69680 Kurdistan_Workers'_Party
[pem_pos=0, final_scores=0.54010516,
final_local_scores=0.51620126, pem=0.0, similarity=19.494188]
11: 69680 Kurdistan_Workers'_Party
[pem_pos=0, final_scores=0.46541554,
final_local_scores=0.43461424, pem=0.0, similarity=17.05947]
12: 854 Anatolia
[pem_pos=0, final_scores=0.3869514,
final_local_scores=0.3830822, pem=-0.23601569,
similarity=16.80514]
13: 69680 Kurdistan_Workers'_Party
[pem_pos=0, final_scores=0.47404653,
final_local_scores=0.44404227, pem=0.0, similarity=17.34082]
1: 17068 Kurdish_people
[pem_pos=1, final_scores=0.07859047,
final_local_scores=-0.0016869903, pem=-0.713598,
similarity=7.920027]
2: 642286 Diyarbakır
[pem_pos=0, final_scores=0.057780072,
final_local_scores=0.003154695, pem=-0.24235578,
similarity=5.5018234]

Sample: b'194_U.S.'
[1 1 U.S. 1] says [2 2 Iraqi Kurds 2] agree ceasefire . [3 3
WASHINGTON 3] 1996-08-23 Leaders of [4 4 Iraq 4] 's two main [
1 1 Kurdish 1] factions agreed on Friday to end six days of
fighting and to attend U.S.-mediated peace talks next month ,
the [5 5 State Department 5] said . Spokesman Glyn Davies
said in a statement that the agreement followed direct [6 6 U.
S. 6] contacts with [7 7 Massoud Barzani 7] , leader of the [8
8 Kurdistan Democratic Party 8] ([9 9 KDP 9]) , and [10 10 J

alal Talabani 10] , leader of the [11 11 Patriotic Union of Kurdistan 11] ([12 12 PUK 12]) . Davies said the two leaders " have agreed to cease the fighting (and) return their forces to the positions held before the current fighting began " on August 17 . He did not give a specific time for the ceasefire but said the [13 13 United States 13] looked forward to " immediate implementation " . The two party leaders had also agreed to meet [14 14 U.S. 14] Assistant Secretary for [15 15 Near Eastern 15] Affairs [16 16 Robert Pelletreau 16] in September " to solidify the cease-fire and to pursue reconciliation " , Davies said . His statement gave no venue or precise date for the meeting . The [17 17 United States 17] has already called on the [2 2 Kurdish 2] factions to hold peace talks in [18 18 London 18] . The [19 19 KDP 19] said Thursday night it had repelled an attack by thousands of [20 20 PUK 20] fighters , killing , wounding or capturing about 400 opposing guerrillas . The fighting has threatened a U.S.-led peace plan to unite the mountainous [21 21 Kurdish 21] region in northern [22 22 Iraq 22] against President [23 23 Saddam Hussein 23] .

1: 3434750 United_States
2: 679693 Iraqi_Kurdistan
3: 108956 Washington,_D.C.
4: 7515928 Iraq
5: 31975 United_States_Department_of_State
6: 3434750 United_States
7: 8122677 Masoud_Barzani
8: 25659771 Kurdistan_Democratic_Party
9: 25659771 Kurdistan_Democratic_Party
10: 440244 Jalal_Talabani
11: 70789 Patriotic_Union_of_Kurdistan
12: 70789 Patriotic_Union_of_Kurdistan
13: 3434750 United_States
14: 3434750 United_States
15: 19323 Middle_East
16: 18299389 Robert_Pelletreau
17: 3434750 United_States
18: 17867 London
19: 25659771 Kurdistan_Democratic_Party
20: 70789 Patriotic_Union_of_Kurdistan
21: 80777 Kurdistan
22: 7515928 Iraq
23: 29490 Saddam_Hussein
1: 679693 Iraqi_Kurdistan
[winner: 80777 Kurdistan, pem_pos=9,
final_scores=0.08815117, final_local_scores=0.031416833,
pem=-2.8001814, similarity=20.255075]
gt: 679693 Iraqi_Kurdistan, pem_pos=6,
final_scores=-0.033907294, final_local_scores=-0.109362006,
pem=-2.7640185, similarity=15.857295]
2: 679693 Iraqi_Kurdistan
[winner: 80777 Kurdistan, pem_pos=9,

```
final_scores=0.03194587, final_local_scores=-0.035870433,
pem=-2.8001814, similarity=18.24709]
    gt: 679693 Iraqi_Kurdistan, pem_pos=6,
final_scores=-0.05725026, final_local_scores=-0.13695937,
pem=-2.7640185, similarity=15.033734 ]
1: 3434750 United_States
    [pem_pos=0, final_scores=0.3992458,
final_local_scores=0.3884378, pem=-0.14075686,
similarity=16.446928]
2: 679693 Iraqi_Kurdistan
    [pem_pos=0, final_scores=0.27979743,
final_local_scores=0.23676124, pem=0.0, similarity=11.155141]
3: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.546985,
final_local_scores=0.56982154, pem=-1.7496481,
similarity=30.60918]
4: 7515928 Iraq
    [pem_pos=0, final_scores=0.6496095,
final_local_scores=0.62431276, pem=-0.25379938,
similarity=24.10065]
5: 31975 United_States_Department_of_State
    [pem_pos=0, final_scores=0.4840514,
final_local_scores=0.49097234, pem=0.0, similarity=18.741306]
6: 3434750 United_States
    [pem_pos=0, final_scores=0.47908878,
final_local_scores=0.47565418, pem=-0.14075686,
similarity=19.04964]
7: 8122677 Masoud_Barzani
    [pem_pos=0, final_scores=0.63342506,
final_local_scores=0.66670537, pem=0.0, similarity=23.98553]
8: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.62820256,
final_local_scores=0.63622063, pem=0.0, similarity=23.075806]
9: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.5793012,
final_local_scores=0.58280325, pem=0.0, similarity=21.481724]
10: 440244 Jalal_Talabani
    [pem_pos=0, final_scores=0.5450315,
final_local_scores=0.5623457, pem=0.0, similarity=20.87123]
11: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.6350134,
final_local_scores=0.65672845, pem=0.0, similarity=23.6878]
12: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.59330225,
final_local_scores=0.61116534, pem=0.0, similarity=22.328106]
13: 3434750 United_States
    [pem_pos=0, final_scores=0.44288045,
final_local_scores=0.4361021, pem=-0.073457584,
similarity=17.503344]
14: 3434750 United_States
    [pem_pos=0, final_scores=0.53976077,
final_local_scores=0.54192907, pem=-0.14075686,
```

```
similarity=21.027414]
15: 19323 Middle_East
    [pem_pos=1, final_scores=0.3594895,
final_local_scores=0.33953553, pem=-0.4631976,
similarity=16.74107]
16: 18299389 Robert_Pelletreau
    [pem_pos=0, final_scores=0.41489917,
final_local_scores=0.5435228, pem=0.0, similarity=20.309517]
17: 3434750 United_States
    [pem_pos=0, final_scores=0.47812343,
final_local_scores=0.47459966, pem=-0.073457584,
similarity=18.652187]
18: 17867 London
    [pem_pos=0, final_scores=0.4254151,
final_local_scores=0.45630717, pem=-0.09086737,
similarity=18.20098]
19: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.5024074,
final_local_scores=0.4988085, pem=0.0, similarity=18.975153]
20: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.53430384,
final_local_scores=0.54671854, pem=0.0, similarity=20.404884]
21: 80777 Kurdistan
    [pem_pos=9, final_scores=0.22213845,
final_local_scores=0.17777759, pem=-2.8001814,
similarity=24.622772]
22: 7515928 Iraq
    [pem_pos=0, final_scores=0.5737931,
final_local_scores=0.54149485, pem=-0.25379938,
similarity=21.6292]
23: 29490 Saddam_Hussein
    [pem_pos=0, final_scores=0.48229766,
final_local_scores=0.46695685, pem=0.0, similarity=18.024637]
1: 80777 Kurdistan
    [pem_pos=9, final_scores=0.08815117,
final_local_scores=0.031416833, pem=-2.8001814,
similarity=20.255075]
2: 80777 Kurdistan
    [pem_pos=9, final_scores=0.03194587,
final_local_scores=-0.035870433, pem=-2.8001814,
similarity=18.24709]
```

Sample: b'195_One'

One teen left dead by attack on [1 1 U.S. 1] slumber party .
[2 2 CHESAPEAKE 2] , Va . 1996-08-23 A knife-wielding neighbour apparently intent on sexual assault invaded a teenage slumber party on Friday , killing one girl and wounding three others , police said . At about 4 a.m. EDT (0800 [3 3 GMT 3]) , a group of teenaged girls were having the overnight party in the Camelot subdivision of this eastern [4 4 Virginia 4] city , when a man entered the house , wielding a knife , threatening to sexually assault

the girls . Detective Richard Black of the Chesapeake Police Department , said a neighbour , Curtis Lee White II , 19 , was arrested in the attack , but had not been charged by late morning on Friday . There were apparently no adults at the party as the father of the family who lived in the house was out of town and the mother died more than a year ago , Black said . The detective said details were sketchy , but two of the teenagers were reportedly downstairs watching television when White allegedly entered the house and told the girls to take off their clothes . He said a male teenager sleeping upstairs reportedly heard the commotion and came downstairs and confronted White , who allegedly stabbed him more than once . The other teenagers also confronted the assailant and three girls , all under 18 , were stabbed , one fatally . " At least two of them were sexually molested , " Black said . He said all of the wounded teenagers were taken to a hospital but none of the injuries were considered life-threatening . Police said the girl who died was identified as Michelle Harper . Her age was not given .

```
1: 3434750 United_States
2: 91342 Chesapeake,_Virginia
3: 12701 Greenwich_Mean_Time
4: 32432 Virginia
1: 3434750 United_States
    [pem_pos=0, final_scores=0.4980808,
final_local_scores=0.50233144, pem=-0.14075686,
similarity=19.845743]
2: 91342 Chesapeake,_Virginia
    [pem_pos=0, final_scores=0.3887776,
final_local_scores=0.41071862, pem=-0.81280684,
similarity=20.766546]
3: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.65541,
final_local_scores=0.7927549, pem=0.0, similarity=27.7471]
4: 32432 Virginia
    [pem_pos=0, final_scores=0.53496116,
final_local_scores=0.5323448, pem=-0.2023721,
similarity=21.076475]
```

Sample: b'196_Glickman'

```
[1 1 Glickman 1] says [2 2 USDA 2] monitoring aflatoxin in [3
3 Texas 3] . [4 4 WASHINGTON 4] 1996-08-23 Agriculture
Secretary [5 5 Dan Glickman 5] said the department was
monitoring reports of aflatoxin found in corn in parts of [6 6
Texas 6] . " We 're always concerned about aflatoxin but we
're on top of it , " [7 7 Glickman 7] told reporters after
addressing a USDA-sponsored farmers ' market . " That 's a
perennial problem . It may be a little more problematic
because of cold , wet conditions but we 're on top of it , "
the secretary said . Asked about reports [8 8 Egypt 8] has
set new levels for a vomitoxin in its purchase of [9 9 U.S. 9]
wheat , the secretary said " I do n't know anything about it
```

" but added that [10 10 USDA 10] officials were " looking at it . "

1: 772726 Dan_Glickman
2: 70896 United_States_Department_of_Agriculture
3: 29810 Texas
4: 108956 Washington,_D.C.
5: 772726 Dan_Glickman
6: 29810 Texas
7: 772726 Dan_Glickman
8: 8087628 Egypt
9: 3434750 United_States
10: 70896 United_States_Department_of_Agriculture
1: 772726 Dan_Glickman
 [pem_pos=0, final_scores=0.5202498,
final_local_scores=0.5297203, pem=0.0, similarity=19.897623]
2: 70896 United_States_Department_of_Agriculture
 [pem_pos=0, final_scores=0.5299363,
final_local_scores=0.5744803, pem=0.0, similarity=21.23335]
3: 29810 Texas
 [pem_pos=0, final_scores=0.75860685,
final_local_scores=0.79770184, pem=-0.16191302,
similarity=28.775234]
4: 108956 Washington,_D.C.
 [pem_pos=0, final_scores=0.5750921,
final_local_scores=0.5855685, pem=-1.7496481,
similarity=31.079103]
5: 772726 Dan_Glickman
 [pem_pos=0, final_scores=0.5864581,
final_local_scores=0.5846756, pem=0.0, similarity=21.5376]
6: 29810 Texas
 [pem_pos=0, final_scores=0.7900323,
final_local_scores=0.8320294, pem=-0.16191302,
similarity=29.799639]
7: 772726 Dan_Glickman
 [pem_pos=0, final_scores=0.4887278,
final_local_scores=0.4952873, pem=0.0, similarity=18.870073]
8: 8087628 Egypt
 [pem_pos=0, final_scores=0.4205798,
final_local_scores=0.5134938, pem=-0.27480945,
similarity=20.907848]
9: 3434750 United_States
 [pem_pos=0, final_scores=0.4758169,
final_local_scores=0.49142283, pem=-0.14075686,
similarity=19.520206]
10: 70896 United_States_Department_of_Agriculture
 [pem_pos=0, final_scores=0.5050044,
final_local_scores=0.547246, pem=0.0, similarity=20.420624]

Sample: b'197_Mass'

[1 1 Mass 1] . governor has trouble winning home support . [2
2 BOSTON 2] 1996-08-23 The 12-year-old daughter of [3 3 Repub
lican 3] Gov . [4 4 William Weld 4] is working to stop his

bid to win a [5 5 U.S. 5] Senate seat because she does n't want to leave [6 6 Massachusetts 6] . [7 7 Weld 7] conceded his 12-year-old daughter , Franny , is " a foot soldier " for [8 8 Democratic 8] incumbent Sen . [9 9 John Kerry 9] , even though she is n't old enough to vote . [10 10 Weld 10] , speaking on [11 11 WBUR-FM 11] radio on Thursday , said he was facing a revolt from his daughter in part because she does not want to leave [12 12 Cambridge 12] , [13 13 Massachusetts 13] , and move to [14 14 Washington 14] . He also said Franny Weld 's best friend , Tracy Roosevelt , might have something to do with her politics . Tracy is the great-granddaughter of [15 15 Democratic 15] former President [16 16 Franklin Roosevelt 16] , and support for [17 17 Democrats 17] runs in the family . The [18 18 Roosevelts 18] are good friends of [19 19 Weld 19] and his wife , [20 20 Susan Roosevelt Weld 20] , a descendant of former President [21 21 Theodore Roosevelt 21] , who won the presidency as a [22 22 Republican 22] .

1: 1645518 Massachusetts
2: 24437894 Boston
3: 32070 Republican_Party_(United_States)
4: 505437 William_Weld
5: 3434750 United_States
6: 1645518 Massachusetts
7: 505437 William_Weld
8: 5043544 Democratic_Party_(United_States)
9: 5122699 John_Kerry
10: 505437 William_Weld
11: 843793 WBUR-FM
12: 5685 Cambridge,_Massachusetts
13: 1645518 Massachusetts
14: 108956 Washington,_D.C.
15: 5043544 Democratic_Party_(United_States)
16: 10979 Franklin_D._Roosevelt
17: 5043544 Democratic_Party_(United_States)
18: 2001979 Roosevelt_family
19: 505437 William_Weld
20: 4406149 Susan_Roosevelt_Weld
21: 30535 Theodore_Roosevelt
22: 32070 Republican_Party_(United_States)
1: 1645518 Massachusetts
 [pem_pos=8, final_scores=0.5479345,
final_local_scores=0.51910275, pem=-3.0163271,
similarity=35.98403]
2: 24437894 Boston
 [pem_pos=0, final_scores=0.6490789,
final_local_scores=0.65781516, pem=-0.23164093,
similarity=24.979929]
3: 32070 Republican_Party_(United_States)
 [pem_pos=0, final_scores=0.50241256,
final_local_scores=0.4716316, pem=-0.1310191,
similarity=18.876642]

4: 505437 William_Weld
[pem_pos=0, final_scores=0.79229516,
final_local_scores=0.7840173, pem=0.0, similarity=27.486353]
5: 3434750 United_States
[pem_pos=0, final_scores=0.53448147,
final_local_scores=0.5597107, pem=-0.14075686,
similarity=21.558052]
6: 1645518 Massachusetts
[pem_pos=0, final_scores=0.6314086,
final_local_scores=0.6102855, pem=-0.1423749,
similarity=23.076107]
7: 505437 William_Weld
[pem_pos=0, final_scores=0.88357764,
final_local_scores=0.8884414, pem=0.0, similarity=30.602575]
8: 5043544 Democratic_Party_(United_States)
[pem_pos=0, final_scores=0.48801428,
final_local_scores=0.45463228, pem=-0.076093495,
similarity=18.070656]
9: 5122699 John_Kerry
[pem_pos=0, final_scores=0.4153816,
final_local_scores=0.40799725, pem=0.0, similarity=16.265165]
10: 505437 William_Weld
[pem_pos=0, final_scores=0.95627266,
final_local_scores=0.96784955, pem=0.0, similarity=32.972275]
11: 843793 WBUR-FM
[pem_pos=0, final_scores=0.36316347,
final_local_scores=0.45167077, pem=0.0, similarity=17.56847]
12: 5685 Cambridge,_Massachusetts
[pem_pos=2, final_scores=0.26338667,
final_local_scores=0.24105665, pem=-1.848646,
similarity=21.336548]
13: 1645518 Massachusetts
[pem_pos=0, final_scores=0.75690776,
final_local_scores=0.7473744, pem=-0.1423749,
similarity=27.167112]
14: 108956 Washington,_D.C.
[pem_pos=0, final_scores=0.5283231,
final_local_scores=0.51611423, pem=-1.7496481,
similarity=29.006449]
15: 5043544 Democratic_Party_(United_States)
[pem_pos=1, final_scores=0.49065536,
final_local_scores=0.45751727, pem=-0.076093495,
similarity=18.15675]
16: 10979 Franklin_D._Roosevelt
[pem_pos=0, final_scores=0.31024206,
final_local_scores=0.2824268, pem=0.0, similarity=12.517892]
17: 5043544 Democratic_Party_(United_States)
[pem_pos=0, final_scores=0.49049604,
final_local_scores=0.45734322, pem=-0.18007085,
similarity=18.717001]
18: 2001979 Roosevelt_family
[pem_pos=0, final_scores=0.30504674,

```
final_local_scores=0.3739556, pem=0.0, similarity=15.249293]
19: 505437 William_Weld
    [pem_pos=0, final_scores=0.90202934,
final_local_scores=0.90859705, pem=0.0, similarity=31.204062]
20: 4406149 Susan_Roosevelt_Weld
    [pem_pos=0, final_scores=0.40630323,
final_local_scores=0.50338686, pem=0.0, similarity=19.11178]
21: 30535 Theodore_Roosevelt
    [pem_pos=0, final_scores=0.42704874,
final_local_scores=0.40660995, pem=0.0, similarity=16.223764]
22: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.4276585,
final_local_scores=0.38997412, pem=-0.1310191,
similarity=16.439821]
```

Sample: b'198_Lufthansa'

```
[1 1 Lufthansa 1] cargo Q2 load factor up 1.7 pct . [2 2 FRA
NKURT 2] 1996-08-23 The following table shows Lufthansa
Cargo AG second quarter 1996 results , based on figures
published by [3 3 Deutsche Lufthansa AG 3] in-house newspaper
Lufthanseat . Available freight-tonne kilometres ( million )
2,389 up 4 pct Revenue freight-tonne kilometres ( million )
1,600 up 7 pct Freight load factor 67.0 up 1.7 pct pts
Revenue from transport ( Dm million ) 820 up 2 pct Revenue
from other services ( Dm million ) 14 down 26 pct Staff costs
( Dm million ) 116 up 8 pct Fuel costs ( Dm million ) 69 up
20 pct Flight-related fees ( Dm million ) 125 up 17 pct - Air
Cargo Newsroom Tel+44 171 542 7706 Fax+44 171 542 5017
```

1: 97733 Lufthansa

2: 10992 Frankfurt

3: 97733 Lufthansa

1: 97733 Lufthansa

```
    [pem_pos=0, final_scores=0.31939077,
final_local_scores=0.25760493, pem=0.0, similarity=11.777158]
```

2: 10992 Frankfurt

```
    [pem_pos=0, final_scores=0.45772523,
final_local_scores=0.51238227, pem=-0.07062824,
similarity=19.76431]
```

3: 97733 Lufthansa

```
    [pem_pos=0, final_scores=0.39785165,
final_local_scores=0.34331152, pem=0.0, similarity=14.334813]
```

Sample: b'199_WSC-India'

```
WSC-India Rice Weather , Aug 23 . SUMMARY- Showers 0.25-1.30
inch ( 6-33 mm ) and locally heavier through much of [1 1 Ind
ia 1] , 75 percent coverage . Isolated showers 0.20-0.70 inch
( 5-18 mm ) in the north . Highs 82-96F ( 28-36C ) . CROP
IMPACT- Conditions remain favorable for the development of
rice in the region . FORECAST- TODAY ... Showers and rain
0.25-1.00 inch ( 6-25 mm ) and locally heavier through most
of central and south central [2 2 India 2] , up to 0.75 inch
( 19 mm ) in 75 percent of north central [3 3 India 3] , and
```

only isolated up to 0.50 inch (13 mm) elsewhere over [4 4 India 4] . Highs 82-96F (28-36C) . TONIGHT ... Variable clouds in southern [5 5 India 5] with showers . Partly cloudy in northern [6 6 India 6] with a few light showers . Lows 68-76F (20-24C) . TOMORROW ... Little change from today 's weather expected . OUTLOOK ... Numerous to scattered showers and thunderstorms in southern and central [7 7 India 7] , and isolated showers to the north Sunday through Tuesday . Temperatures near normal . Source : Weather Services Corporation

```
1: 14533 India
2: 14533 India
3: 14533 India
4: 14533 India
5: 14533 India
6: 14533 India
7: 14533 India
1: 14533 India
    [pem_pos=0, final_scores=0.59497935,
final_local_scores=0.5218332, pem=-0.13622136,
similarity=20.40305]
2: 14533 India
    [pem_pos=0, final_scores=0.53370315,
final_local_scores=0.45489824, pem=-0.13622136,
similarity=18.405577]
3: 14533 India
    [pem_pos=0, final_scores=0.58028895,
final_local_scores=0.5057861, pem=-0.13622136,
similarity=19.924173]
4: 14533 India
    [pem_pos=0, final_scores=0.6013735,
final_local_scores=0.52881783, pem=-0.13622136,
similarity=20.611485]
5: 14533 India
    [pem_pos=0, final_scores=0.6249114,
final_local_scores=0.5545294, pem=-0.13622136,
similarity=21.37877]
6: 14533 India
    [pem_pos=0, final_scores=0.5933469,
final_local_scores=0.5200499, pem=-0.13622136,
similarity=20.349833]
7: 14533 India
    [pem_pos=0, final_scores=0.53588665,
final_local_scores=0.45728338, pem=-0.13622136,
similarity=18.476753]
```

Sample: b'200_Washington'

```
[1 1 Washington 1] to curb [2 2 Tamil 2] support in [3 3 U.S.
3] - [4 4 Sri Lanka 4] . [5 5 COLOMBO 5] 1996-08-23 [6 6 Sri L
anka 6] said on Friday the [7 7 United States 7] had promised
to stamp out any illegal activities on [8 8 U.S. 8] soil
directed against the island 's government . The [9 9 Sri Lank
```

an 9] foreign ministry said in a statement : " The [10 10 United States 10] government sympathised with the current predicament [11 11 Sri Lanka 11] was facing . " The statement said the [12 12 U.S. 12] government " would do all within its prevailing legal framework to prevent the use of [13 13 American 13] soil to perpetrate violence against the democratic government of [14 14 Sri Lanka 14] " . It said the [15 15 U.S. State Department 15] 's coordinator for counter terrorism , Philip Wilcox , had expressed [16 16 Washington 16] 's support for the government when he visited [17 17 Colombo 17] this week . [18 18 Colombo 18] has said it believes [19 19 Tamil 19] rebels , fighting a 13-year war for independence against the government , finance their military activity through funds extorted from expatriate [20 20 Sri Lankans 20] in western countries such as the [21 21 United States 21] . [2 2 22 U.S. 22] embassy officials in [23 23 Colombo 23] were not immediately available to comment on the report . [24 24 Colombo 24] estimates more than 50,000 people have been killed in the war between government forces and the [25 25 Liberation Tigers of Tamil Eelam 25] rebels in the island 's north and east .

1: 108956 Washington,_D.C.
2: 4096004 Sri_Lankan_Tamil_people
3: 3434750 United_States
4: 26750 Sri_Lanka
5: 56636 Colombo
6: 26750 Sri_Lanka
7: 3434750 United_States
8: 3434750 United_States
9: 26750 Sri_Lanka
10: 3434750 United_States
11: 26750 Sri_Lanka
12: 3434750 United_States
13: 3434750 United_States
14: 26750 Sri_Lanka
15: 31975 United_States_Department_of_State
16: 108956 Washington,_D.C.
17: 56636 Colombo
18: 56636 Colombo
19: 4096004 Sri_Lankan_Tamil_people
20: 26750 Sri_Lanka
21: 3434750 United_States
22: 3434750 United_States
23: 56636 Colombo
24: 56636 Colombo
25: 18606 Liberation_Tigers_of_Tamil_Eelam
1: 108956 Washington,_D.C.
[pem_pos=0, final_scores=0.37825143,
final_local_scores=0.36427945, pem=-1.7496481,
similarity=24.475397]
2: 4096004 Sri_Lankan_Tamil_people
[pem_pos=4, final_scores=0.29867393,

```
final_local_scores=0.30805862, pem=-2.8080099,  
similarity=28.55319]  
3: 3434750 United_States  
    [pem_pos=0, final_scores=0.5486651,  
final_local_scores=0.5198247, pem=-0.14075686,  
similarity=20.367777]  
4: 26750 Sri_Lanka  
    [pem_pos=0, final_scores=0.6094591,  
final_local_scores=0.57863224, pem=-0.3027737,  
similarity=23.003784]  
5: 56636 Colombo  
    [pem_pos=0, final_scores=0.57069105,  
final_local_scores=0.5530913, pem=-0.19393882,  
similarity=21.64973]  
6: 26750 Sri_Lanka  
    [pem_pos=0, final_scores=0.54814655,  
final_local_scores=0.51165754, pem=-0.3027737,  
similarity=21.005125]  
7: 3434750 United_States  
    [pem_pos=0, final_scores=0.50929886,  
final_local_scores=0.47682303, pem=-0.073457584,  
similarity=18.718538]  
8: 3434750 United_States  
    [pem_pos=0, final_scores=0.5178812,  
final_local_scores=0.486198, pem=-0.14075686,  
similarity=19.364288]  
9: 26750 Sri_Lanka  
    [pem_pos=0, final_scores=0.56871146,  
final_local_scores=0.53412163, pem=-0.27403015,  
similarity=21.519186]  
10: 3434750 United_States  
    [pem_pos=0, final_scores=0.50452673,  
final_local_scores=0.47161025, pem=-0.073457584,  
similarity=18.562977]  
11: 26750 Sri_Lanka  
    [pem_pos=0, final_scores=0.4363187,  
final_local_scores=0.38950264, pem=-0.3027737,  
similarity=17.35978]  
12: 3434750 United_States  
    [pem_pos=0, final_scores=0.5393738,  
final_local_scores=0.50967544, pem=-0.14075686,  
similarity=20.064901]  
13: 3434750 United_States  
    [pem_pos=0, final_scores=0.5453854,  
final_local_scores=0.51624215, pem=-0.12994486,  
similarity=20.20207]  
14: 26750 Sri_Lanka  
    [pem_pos=0, final_scores=0.51050234,  
final_local_scores=0.470537, pem=-0.3027737,  
similarity=19.778006]  
15: 31975 United_States_Department_of_State  
    [pem_pos=0, final_scores=0.37122285,
```

```
final_local_scores=0.3891222, pem=0.0, similarity=15.701895]
16: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.31971103,
final_local_scores=0.30033296, pem=-1.7496481,
similarity=22.567106]
17: 56636 Colombo
    [pem_pos=0, final_scores=0.47127396,
final_local_scores=0.4444933, pem=-0.19393882,
similarity=18.408949]
18: 56636 Colombo
    [pem_pos=0, final_scores=0.46179444,
final_local_scores=0.43413836, pem=-0.19393882,
similarity=18.099937]
19: 4096004 Sri_Lankan_Tamil_people
    [pem_pos=4, final_scores=0.3325886,
final_local_scores=0.34510517, pem=-2.8080099,
similarity=29.658733]
20: 26750 Sri_Lanka
    [pem_pos=1, final_scores=0.40434206,
final_local_scores=0.354573, pem=-0.5013852,
similarity=17.39749]
21: 3434750 United_States
    [pem_pos=0, final_scores=0.5223568,
final_local_scores=0.4910869, pem=-0.073457584,
similarity=19.1442]
22: 3434750 United_States
    [pem_pos=0, final_scores=0.5514154,
final_local_scores=0.522829, pem=-0.14075686,
similarity=20.457432]
23: 56636 Colombo
    [pem_pos=0, final_scores=0.48506373,
final_local_scores=0.45955652, pem=-0.19393882,
similarity=18.858465]
24: 56636 Colombo
    [pem_pos=0, final_scores=0.5194494,
final_local_scores=0.49711758, pem=-0.19393882,
similarity=19.979362]
25: 18606 Liberation_Tigers_of_Tamil_Eelam
    [pem_pos=0, final_scores=0.52570355,
final_local_scores=0.5840856, pem=0.0, similarity=21.51999]
```

Sample: b'201_Nepal'

[1 1 Nepal 1] 's king leaves on week-long visit to [2 2 China 2] . [3 3 KATHMANDU 3] 1996-08-23 King [4 4 Birendra 4] left [5 5 Nepal 5] on Friday for a week-long visit to [6 6 China 6] , his eighth since ascending the throne in 1972 , officials said . The constitutional monarch , who last visited [7 7 China 7] in 1993 , was scheduled to meet [8 8 Chinese 8] President [9 9 Jiang Zemin 9] and Premier [10 10 Li Peng 10] during his visit , they said . Foreign ministry officials gave no details of the issues the king , who was accompanied by Foreign Minister [11 11 Prakash Chandra Lohani 11] , would

discuss with [12 12 Chinese 12] leaders . The [13 13 Himalayan 13] kingdom , sandwiched between [14 14 China 14] and [15 15 India 15] , has traditionally sought to maintain close cooperation with its giant neighbours , and an equal distance from the two . The 50-year old monarch was accompanied by Queen Aishwarya on a flight to the [16 16 Tibetan 16] capital of [17 17 Lhasa 17] . The king will visit [18 18 Chongqing 18] before arriving in the [19 19 Chinese 19] capital , [20 20 Beijing 20] , early next week , officials said .

1: 171166 Nepal
2: 5405 China
3: 17168 Kathmandu
4: 313169 Birendra_of_Nepal
5: 171166 Nepal
6: 5405 China
7: 5405 China
8: 5405 China
9: 95871 Jiang_Zemin
10: 44533 Li_Peng
11: 15532150 Prakash_Chandra_Lohani
12: 5405 China
13: 37966 Himalayas
14: 5405 China
15: 14533 India
16: 31516 Tibet
17: 238741 Lhasa
18: 171752 Chongqing
19: 5405 China
20: 18603746 Beijing
1: 171166 Nepal
[pem_pos=0, final_scores=0.68429345,
final_local_scores=0.6679555, pem=-0.22887385,
similarity=25.267488]
2: 5405 China
[pem_pos=0, final_scores=0.5731742,
final_local_scores=0.5214195, pem=-0.15669827,
similarity=20.502062]
3: 17168 Kathmandu
[pem_pos=0, final_scores=0.39057773,
final_local_scores=0.38733214, pem=0.0, similarity=15.648477]
4: 313169 Birendra_of_Nepal
[pem_pos=0, final_scores=0.4302835,
final_local_scores=0.49830467, pem=0.0, similarity=18.960117]
5: 171166 Nepal
[pem_pos=0, final_scores=0.61141235,
final_local_scores=0.588344, pem=-0.22887385,
similarity=22.891722]
6: 5405 China
[pem_pos=0, final_scores=0.52478313,
final_local_scores=0.46855962, pem=-0.15669827,
similarity=18.924616]
7: 5405 China

```
[pem_pos=0, final_scores=0.51444453,
final_local_scores=0.45726627, pem=-0.15669827,
similarity=18.5876]
8: 5405 China
    [pem_pos=1, final_scores=0.75659806,
final_local_scores=0.72178227, pem=-1.4196125,
similarity=33.34921]
9: 95871 Jiang_Zemin
    [pem_pos=0, final_scores=0.5450073,
final_local_scores=0.59923524, pem=0.0, similarity=21.972088]
10: 44533 Li_Peng
    [pem_pos=0, final_scores=0.52976567,
final_local_scores=0.583297, pem=0.0, similarity=21.49646]
11: 15532150 Prakash_Chandra_Lohani
    [pem_pos=0, final_scores=0.5183907,
final_local_scores=0.6965293, pem=0.0, similarity=24.875538]
12: 5405 China
    [pem_pos=1, final_scores=0.6863609,
final_local_scores=0.6450588, pem=-1.4196125,
similarity=31.059628]
13: 37966 Himalayas
    [pem_pos=0, final_scores=0.37800258,
final_local_scores=0.38499558, pem=0.0, similarity=15.57875]
14: 5405 China
    [pem_pos=0, final_scores=0.581011,
final_local_scores=0.52998006, pem=-0.15669827,
similarity=20.757524]
15: 14533 India
    [pem_pos=0, final_scores=0.5852404,
final_local_scores=0.5715835, pem=-0.13622136,
similarity=21.887697]
16: 31516 Tibet
    [pem_pos=1, final_scores=0.5192837,
final_local_scores=0.4926861, pem=-1.2825032,
similarity=25.766901]
17: 238741 Lhasa
    [pem_pos=0, final_scores=0.37250555,
final_local_scores=0.36109066, pem=0.0, similarity=14.865378]
18: 171752 Chongqing
    [pem_pos=0, final_scores=0.3463912,
final_local_scores=0.33349884, pem=0.0, similarity=14.041984]
19: 5405 China
    [pem_pos=1, final_scores=0.71388966,
final_local_scores=0.67512983, pem=-1.4196125,
similarity=31.957008]
20: 18603746 Beijing
    [pem_pos=0, final_scores=0.43675727,
final_local_scores=0.43650973, pem=-0.042161036,
similarity=17.345312]
```

Sample: b'202_Nepal'

[1 1 Nepal 1] man held for keeping child servant in chains .

[2 2 KATHMANDU 2] 1996-08-23 [3 3 Nepali 3] police said on Friday they arrested a man who allegedly kept a child servant bound in chains so that he would not run away when his employer was out to work . Madhusudan Munakarmi was arrested on Thursday after his neighbours informed police about the plight of 12-year old Dhiraj K.C. , who told police his employer used to tie him up with iron chains and locks concealed under his clothes . The neighbours in [4 4 Kathmandu 4] called the police when they saw Dheeraj , employed by the man for the past nine months , limping because of the chains . " I feared he would flee from work or steal my belongings , " the Kathmandu Post newspaper quoted Munakarmi as saying after his arrest . If convicted , he faces a maximum of three years in jail under [5 5 Nepal 5] 's child protection laws .

1: 171166 Nepal
2: 17168 Kathmandu
3: 171166 Nepal
4: 17168 Kathmandu
5: 171166 Nepal
1: 171166 Nepal
[pem_pos=0, final_scores=0.6617972,
final_local_scores=0.6133514, pem=-0.22887385,
similarity=23.637993]
2: 17168 Kathmandu
[pem_pos=0, final_scores=0.44441772,
final_local_scores=0.39930326, pem=0.0, similarity=16.005718]
3: 171166 Nepal
[pem_pos=1, final_scores=0.59188396,
final_local_scores=0.5369818, pem=-1.4646683,
similarity=28.079414]
4: 17168 Kathmandu
[pem_pos=0, final_scores=0.4032309,
final_local_scores=0.35431293, pem=0.0, similarity=14.663117]
5: 171166 Nepal
[pem_pos=0, final_scores=0.70340437,
final_local_scores=0.65880096, pem=-0.22887385,
similarity=24.994299]

Sample: b'203_OPTIONS'
OPTIONS - Euro debt vols seen regrouping after fall . [1 1 LO NDON 1] 1996-08-23 Implied volatility of [2 2 European 2] bond and interest rate options should stabilise around current levels until early next week after falling before and after this week 's German-led cut in interest rates , traders said . " Volatility has come off a lot . We 're looking for it to stabilise now , " said one Euromark options trader at a [3 3 U.S. 3] bank . A trader at a [4 4 Japanese 4] bank said Euromark volatility now stood at 14.00 for September contract , 16.75 for December , 19.50 for March and 21.25 for June . This compared with midweek levels , before the welter of interest rate cuts , of 18.50 for September , 20.00 for

December , 22.00 for March and 23.5 for June , he said . At 1347 [5 5 GMT 5] , December Euromark futures were trading at 96.78 , two basis points down on the day . He said the sell-off in June vols might have been overdone , which could offer value at current levels . He said caps and floors would be well bid after the round of interest rate cuts due to the fact these rates should stay low at the short end . The size of the [6 6 Bundesbank 6] 's repo rate cut , to 3.00 percent from 3.30 percent , took markets by surprise . " Volatility has a bid to it -- longer-dated volatility more than short-dated because the event people were buying for has passed and now perhaps people will sell some short dated vol and buy some long dated vol , " he said . " Long dated volatility has been low this year , so it is still at levels which are not historically high . " It is not a dangerous level to own vol. You are not going to lose a lot and you could make quite a bit . " He said volatility levels should be stable until markets reassess the situation after a long weekend in [7 7 Britain 7] . Paribas Capital Markets OTC options specialist Robert Coughlan said that if volatility continued lower for the rest of Friday in over-the-counter 10-year Bunds , it should be higher next Tuesday . He said the absence on holiday of many market makers was a main factor behind falls this week in volatility in high-yielding markets such as [8 8 Italy 8] , [9 9 Spain 9] and [10 10 Sweden 10] . Coughlan said the market had more downside than upside potential , but a fall was not likely to be of significant size . " I recommend people sell strangles in a number of markets -- in [11 11 Germany 11] and [12 12 France 12] in particular . " With high-yielding markets [13 13 Italy 13] will be a lot more vulnerable in September on economic and political fronts , so I would use current the low level of vol to buy options . " So sell options on Bunds and [14 14 France 14] to enhance yield and buy options on [15 15 Italy 15] , " Coughlan said .. -- Stephen Nisbet , International Bonds +44 171 542 6320

1: 17867 London
2: 9317 European_Union
3: 3434750 United_States
4: 15573 Japan
5: 12701 Greenwich_Mean_Time
6: 274294 Deutsche_Bundesbank
7: 31717 United_Kingdom
8: 14532 Italy
9: 26667 Spain
10: 5058739 Sweden
11: 11867 Germany
12: 5843419 France
13: 14532 Italy
14: 5843419 France
15: 14532 Italy
1: 17867 London

```
[pem_pos=0, final_scores=0.67099386,
final_local_scores=0.67864, pem=-0.09086737,
similarity=24.835835]
2: 9317 European_Union
    [pem_pos=1, final_scores=0.14217664,
final_local_scores=0.12879139, pem=-2.9373715,
similarity=23.906988]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.4393304,
final_local_scores=0.4053076, pem=-0.14075686,
similarity=16.950357]
4: 15573 Japan
    [pem_pos=0, final_scores=0.79846126,
final_local_scores=0.8243156, pem=-0.99076414,
similarity=34.076862]
5: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.64414495,
final_local_scores=0.74597913, pem=0.0, similarity=26.351217]
6: 274294 Deutsche_Bundesbank
    [pem_pos=0, final_scores=0.2686612,
final_local_scores=0.35846686, pem=0.0, similarity=14.787079]
7: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5000678,
final_local_scores=0.46428722, pem=-0.61114657,
similarity=21.268478]
8: 14532 Italy
    [pem_pos=0, final_scores=0.7201241,
final_local_scores=0.70349663, pem=-0.20612453,
similarity=26.204393]
9: 26667 Spain
    [pem_pos=0, final_scores=0.6751837,
final_local_scores=0.6667802, pem=-0.18857051,
similarity=25.013239]
10: 5058739 Sweden
    [pem_pos=0, final_scores=0.64947575,
final_local_scores=0.6499054, pem=-0.2135508,
similarity=24.645508]
11: 11867 Germany
    [pem_pos=0, final_scores=0.66192824,
final_local_scores=0.63968414, pem=-0.15529114,
similarity=24.02366]
12: 5843419 France
    [pem_pos=0, final_scores=0.6803667,
final_local_scores=0.65332013, pem=-0.16240834,
similarity=24.46929]
13: 14532 Italy
    [pem_pos=0, final_scores=0.6993525,
final_local_scores=0.6808068, pem=-0.20612453,
similarity=25.527285]
14: 5843419 France
    [pem_pos=0, final_scores=0.6144842,
final_local_scores=0.5813535, pem=-0.16240834,
```

```
similarity=22.321661]
15: 14532 Italy
    [pem_pos=0, final_scores=0.61731106,
final_local_scores=0.59118897, pem=-0.20612453,
similarity=22.85291]

Sample: b'204_Goldman'
[1 1 Goldman Sachs 1] sets warrants on [2 2 Continental 2] .
[3 3 LONDON 3] 1996-08-23 Goldman Sachs & Co Wertpapier GmbH
has issued a total of five million American-style call
warrants , on [4 4 Continental AG 4] , lead manager [5 5 Gold
man Sachs & Co 5] said . One warrant controls one share .
STRIKE PRICE 25.00 [6 6 DEM 6] PREMIUM 10.12 PCT ISSUE PRICE
2.42 [7 7 DEM 7] GEARING 10.29 X EXERCISE PERIOD 02.SEP.
96-21.NOV.97 PAYDATE 30.AUG.96 LISTING DDF FFT STG MIN EXER
LOT 100 SPOT REFERENCE 24.90 [8 8 DEM 8] -- Reuter London
Newsroom +44 171 542 7658
1: 335244 Goldman_Sachs
2: 1604474 Continental_AG
3: 17867 London
4: 1604474 Continental_AG
5: 335244 Goldman_Sachs
6: 88160 Deutsche_Mark
7: 88160 Deutsche_Mark
8: 88160 Deutsche_Mark
1: 335244 Goldman_Sachs
    [pem_pos=0, final_scores=0.4488334,
final_local_scores=0.48687315, pem=0.0, similarity=18.618979]
2: 1604474 Continental_AG
    [pem_pos=2, final_scores=0.29180676,
final_local_scores=0.32086065, pem=-2.1237617,
similarity=25.214182]
3: 17867 London
    [pem_pos=0, final_scores=0.61190003,
final_local_scores=0.72104275, pem=-0.09086737,
similarity=26.101217]
4: 1604474 Continental_AG
    [pem_pos=0, final_scores=0.41228414,
final_local_scores=0.45246392, pem=0.0, similarity=17.592138]
5: 335244 Goldman_Sachs
    [pem_pos=0, final_scores=0.5470892,
final_local_scores=0.5942027, pem=0.0, similarity=21.821907]
6: 88160 Deutsche_Mark
    [pem_pos=2, final_scores=0.62674147,
final_local_scores=0.64045614, pem=-1.9186858,
similarity=33.636314]
7: 88160 Deutsche_Mark
    [pem_pos=2, final_scores=0.54477924,
final_local_scores=0.55092484, pem=-1.9186858,
similarity=30.96452]
8: 88160 Deutsche_Mark
    [pem_pos=2, final_scores=0.6890529,
```

```
final_local_scores=0.7085219, pem=-1.9186858,  
similarity=35.66753]
```

Sample: b'205_Legal'

Legal challenge to [1 1 Diana 1] delayed by jail term . [2 2 LONDON 2] 1996-08-23 A [3 3 British 3] photographer branded a stalker by Princess [4 4 Diana 4] has been forced to postpone a legal challenge to a ban on approaching her because he 's been jailed for criminal damage , his lawyer said on Friday . Martin Stenning started a 12-week jail sentence on Thursday just as he was preparing to contest an injunction obtained by [5 5 Diana 5] banning him from coming within 300 metres (yards) of her . " We were in the process of preparing a detailed affidavit responding to the Princess 's affidavit and expected to go to court in the next couple of weeks , " said Stenning 's lawyer , Benedict Birnberg . " But everything has been put on ice now . " Birnberg told [6 6 Reuters 6] that the challenge to the injunction would be delayed until Stenning was released . Stenning threw a brick through the window of a van in February after an argument with a driver when he was working as a motorcycle dispatch rider . Stenning , who has previous convictions , is expected to appeal against the sentence . Magistrates also ordered him to pay compensation of 182 pounds (\$ 282) . The freelance photographer was branded a stalker by [7 7 Diana 7] , whose divorce from heir-to-the-throne Prince [8 8 Charles 8] is due to become final next week , after persistently trailing her on his motorcycle . In an affidavit , the princess said that in chasing her Stenning had got so close that he twice smashed into her car and pushed her when she tried to remove the film from his camera . Stenning has rejected [9 9 Diana 9] 's claims and said he was being made a scapegoat to scare off press photographers .

```
1: 24095 Diana,_Princess_of_Wales  
2: 17867 London  
3: 31717 United_Kingdom  
4: 24095 Diana,_Princess_of_Wales  
5: 24095 Diana,_Princess_of_Wales  
6: 18998750 Reuters  
7: 24095 Diana,_Princess_of_Wales  
8: 125248 Charles,_Prince_of_Wales  
9: 24095 Diana,_Princess_of_Wales  
1: 24095 Diana,_Princess_of_Wales  
    [pem_pos=0, final_scores=1.1726345,  
final_local_scores=1.1624696, pem=0.0, similarity=38.780125]  
2: 17867 London  
    [pem_pos=0, final_scores=0.6171301,  
final_local_scores=0.64823496, pem=-0.09086737,  
similarity=23.928488]  
3: 31717 United_Kingdom  
    [pem_pos=0, final_scores=0.46157187,  
final_local_scores=0.49598426, pem=-0.48936236,
```

```
similarity=21.552101]
4: 24095 Diana,_Princess_of_Wales
    [pem_pos=2, final_scores=0.691039,
final_local_scores=0.6347411, pem=-2.4220667,
similarity=36.203228]
5: 24095 Diana,_Princess_of_Wales
    [pem_pos=0, final_scores=1.1742501,
final_local_scores=1.1642344, pem=0.0, similarity=38.832787]
6: 18998750 Reuters
    [pem_pos=0, final_scores=0.3423413,
final_local_scores=0.3934486, pem=0.0, similarity=15.831003]
7: 24095 Diana,_Princess_of_Wales
    [pem_pos=0, final_scores=1.0713124,
final_local_scores=1.0517907, pem=0.0, similarity=35.47725]
8: 125248 Charles,_Prince_of_Wales
    [pem_pos=2, final_scores=0.51926404,
final_local_scores=0.48696464, pem=-2.8139348,
similarity=33.924324]
9: 24095 Diana,_Princess_of_Wales
    [pem_pos=0, final_scores=1.1688278,
final_local_scores=1.1583114, pem=0.0, similarity=38.656033]
```

Sample: b'206_Jordan'

[1 1 Jordan 1] expels [2 2 Iraqi 2] diplomat . [3 3 AMMAN 3] 1996-08-23 [4 4 Jordan 4] has asked an [5 5 Iraqi 5] diplomat to leave the kingdom for carrying out duties incompatible with diplomatic norms , an official source said on Friday . The move came after [6 6 Amman 6] blamed [7 7 Iraq 7] and a pro-Baghdad local political party for last week 's worst unrest in seven years after a government decision to double prices of bread . The government declined comment . " [8 8 Jordan 8] has asked Mr. Adel Ibrahim , the [9 9 Iraqi 9] embassy 's press attache , to leave because he was carrying out duties incompatible with diplomatic norms , " the source told [10 10 Reuters 10] . He said Ibrahim was still in [11 11 Amman 11] . The Jordanian Arab Socialist Baath Party has denied involvement in unrest which it blamed on government policies and rising economic hardship . The riots , which shook [12 12 Jordan 12] for two days , broke out after last Friday 's main prayers in the southern town of [13 13 Karak 13] and spread to [14 14 Amman 14] .

1: 7515964 Jordan
2: 7515928 Iraq
3: 45007 Amman
4: 7515964 Jordan
5: 7515928 Iraq
6: 45007 Amman
7: 7515928 Iraq
8: 7515964 Jordan
9: 7515928 Iraq
10: 18998750 Reuters
11: 45007 Amman

12: 7515964 Jordan
13: 307576 Al_Karak
14: 45007 Amman
1: 7515964 Jordan
[pem_pos=0, final_scores=0.6141589,
final_local_scores=0.5603197, pem=-0.53647494,
similarity=23.728205]
2: 7515928 Iraq
[pem_pos=0, final_scores=0.5354817,
final_local_scores=0.48453248, pem=-0.48469657,
similarity=21.184984]
3: 45007 Amman
[pem_pos=0, final_scores=0.575621,
final_local_scores=0.5352977, pem=0.0, similarity=20.064064]
4: 7515964 Jordan
[pem_pos=0, final_scores=0.59655964,
final_local_scores=0.54109514, pem=-0.53647494,
similarity=23.154505]
5: 7515928 Iraq
[pem_pos=0, final_scores=0.5225968,
final_local_scores=0.4704576, pem=-0.48469657,
similarity=20.764961]
6: 45007 Amman
[pem_pos=0, final_scores=0.4302184,
final_local_scores=0.3764674, pem=0.0, similarity=15.324251]
7: 7515928 Iraq
[pem_pos=0, final_scores=0.7032839,
final_local_scores=0.6678309, pem=-0.25379938,
similarity=25.399319]
8: 7515964 Jordan
[pem_pos=0, final_scores=0.5614905,
final_local_scores=0.5027874, pem=-0.53647494,
similarity=22.011326]
9: 7515928 Iraq
[pem_pos=0, final_scores=0.5779475,
final_local_scores=0.5309199, pem=-0.48469657,
similarity=22.569277]
10: 18998750 Reuters
[pem_pos=0, final_scores=0.4052288,
final_local_scores=0.44321632, pem=0.0, similarity=17.316172]
11: 45007 Amman
[pem_pos=0, final_scores=0.42943102,
final_local_scores=0.3756073, pem=0.0, similarity=15.298584]
12: 7515964 Jordan
[pem_pos=0, final_scores=0.5663219,
final_local_scores=0.50806504, pem=-0.53647494,
similarity=22.168821]
13: 307576 Al_Karak
[pem_pos=0, final_scores=0.5060933,
final_local_scores=0.5379639, pem=-0.6855135,
similarity=23.87156]
14: 45007 Amman

```
[pem_pos=0, final_scores=0.44442546,
final_local_scores=0.3919865, pem=0.0, similarity=15.787372]
```

Sample: b'207_Kurd'

Kurd rebels to free Turkish soldier prisoners . [1 1 DOHUK 1] , [2 2 Iraq 2] 1996-08-23 Turkish Kurd guerrillas said on Friday they would free seven Turkish soldiers they hold in northern [3 3 Iraq 3] under a tentative [4 4 Islamist 4] peace bid . " ...For the sake of safety we are asking for their family members or the authorities to come and pick them up , " [5 5 Kurdistan Workers Party 5] ([6 6 PKK 6]) central committee member Riza Altun told journalists near the [7 7 Iraqi 7] city of [8 8 Dohuk 8] . [9 9 PKK 9] guerrillas would accompany the soldiers , captured last spring in one of Turkey 's frequent cross-border drives , until they could be handed over , he said . Their release has been negotiated by [10 10 Islamist 10] writer Ismail Nacar as part of a wider effort , partly backed by Prime Minister [11 11 Necmettin Erbakan 11] , to find a political solution to Turkey 's [12 12 Kurdish 12] problem . [13 13 Erbakan 13] has encouraged Nacar 's bid but has ruled out direct talks with the rebels . The [14 14 PKK 14] often uses bases in northern [15 15 Iraq 15] in its fight for autonomy or independence in southeast Turkey . More than 20,000 people have died in 12 years of fighting between the guerrillas and Turkish forces .

1: 566384 Dohuk
2: 7515928 Iraq
3: 7515928 Iraq
4: 6037917 Islam
5: 69680 Kurdistan_Workers'_Party
6: 69680 Kurdistan_Workers'_Party
7: 7515928 Iraq
8: 566384 Dohuk
9: 69680 Kurdistan_Workers'_Party
10: 6037917 Islam
11: 1328699 Necmettin_Erbakan
12: 80777 Kurdistan
13: 1328699 Necmettin_Erbakan
14: 69680 Kurdistan_Workers'_Party
15: 7515928 Iraq
1: 566384 Dohuk
[pem_pos=0, final_scores=0.33522302,
final_local_scores=0.37130582, pem=0.0, similarity=15.170219]
2: 7515928 Iraq
[pem_pos=0, final_scores=0.61220086,
final_local_scores=0.5796551, pem=-0.25379938,
similarity=22.767979]
3: 7515928 Iraq
[pem_pos=0, final_scores=0.6160974,
final_local_scores=0.5839115, pem=-0.25379938,
similarity=22.894997]
4: 6037917 Islam

```
[pem_pos=1, final_scores=0.3238445,
final_local_scores=0.32758752, pem=-1.3455799,
similarity=21.18305]
5: 69680 Kurdistan_Workers'_Party
    [pem_pos=0, final_scores=0.6497353,
final_local_scores=0.64311135, pem=0.0, similarity=23.281439]
6: 69680 Kurdistan_Workers'_Party
    [pem_pos=0, final_scores=0.53100747,
final_local_scores=0.51341933, pem=0.0, similarity=19.41117]
7: 7515928 Iraq
    [pem_pos=0, final_scores=0.54365474,
final_local_scores=0.5047789, pem=-0.48469657,
similarity=21.789179]
8: 566384 Dohuk
    [pem_pos=0, final_scores=0.29108852,
final_local_scores=0.32309568, pem=0.0, similarity=13.731532]
9: 69680 Kurdistan_Workers'_Party
    [pem_pos=0, final_scores=0.45999616,
final_local_scores=0.4358502, pem=0.0, similarity=17.096352]
10: 6037917 Islam
    [pem_pos=1, final_scores=0.36378127,
final_local_scores=0.37121242, pem=-1.3455799,
similarity=22.484903]
11: 1328699 Necmettin_Erbakan
    [pem_pos=0, final_scores=0.5680588,
final_local_scores=0.6341694, pem=0.0, similarity=23.014593]
12: 80777 Kurdistan
    [pem_pos=9, final_scores=0.33171672,
final_local_scores=0.2879225, pem=-2.8001814,
similarity=27.909716]
13: 1328699 Necmettin_Erbakan
    [pem_pos=0, final_scores=0.3627053,
final_local_scores=0.40985197, pem=0.0, similarity=16.320513]
14: 69680 Kurdistan_Workers'_Party
    [pem_pos=0, final_scores=0.49571306,
final_local_scores=0.4748655, pem=0.0, similarity=18.260647]
15: 7515928 Iraq
    [pem_pos=0, final_scores=0.5864446,
final_local_scores=0.55152035, pem=-0.25379938,
similarity=21.928381]
```

Sample: b'208_SOLIDERE'

[1 1 SOLIDERE 1] shares mixed on market . [2 2 BEIRUT 2]
1996-08-23 [3 3 SOLIDERE 3] shares were mixed on Friday on
the privately-operated Beirut Secondary Market (BSM) . A
shares -- distributed to former holders of property rights in
the [4 4 Beirut 4] central district [5 5 SOLIDERE 5] is
rebuilding -- closed at \$ 104.625 unchanged from Thursday . B
shares -- issued in a \$ 650-million subscription in January
1994 -- rose to \$ 106.5 from \$ 106.375 a day earlier .
Turnover on BSM , which trades only [6 6 SOLIDERE 6] shares ,
was 8,049 shares from Thursday 's 8,757 and value was \$

850,968 from \$ 918,288 . On the official [7 7 Beirut Stock Exchange 7] , only 1,185 Ciments Libanais shares were traded at \$ 1.1875 compared with 2,036 shares traded on Thursday at the same price . There was no trade in any of the three other listed companies : Ciments Blancs , [8 8 Eternit 8] and Uniceramic . The BLOM Stock Index which covers both markets rose 0.04 percent to 903.09 and the LISPI index rose 0.02 percent to 81.58 . - [9 9 Beirut 9] editorial (961 1)
864148 353078 861723
1: 2886010 Solidere
2: 37428 Beirut
3: 2886010 Solidere
4: 37428 Beirut
5: 2886010 Solidere
6: 2886010 Solidere
7: 9848049 Beirut_Stock_Exchange
8: 27417693 Eternit
9: 37428 Beirut
1: 2886010 Solidere
[pem_pos=0, final_scores=0.5601157,
final_local_scores=0.53533083, pem=0.0, similarity=20.065052]
2: 37428 Beirut
[pem_pos=0, final_scores=0.48005944,
final_local_scores=0.47885478, pem=-0.046614014,
similarity=18.633188]
3: 2886010 Solidere
[pem_pos=0, final_scores=0.5510909,
final_local_scores=0.5254726, pem=0.0, similarity=19.770863]
4: 37428 Beirut
[pem_pos=0, final_scores=0.40076917,
final_local_scores=0.39224225, pem=-0.046614014,
similarity=16.048498]
5: 2886010 Solidere
[pem_pos=0, final_scores=0.59209996,
final_local_scores=0.57026875, pem=0.0, similarity=21.10767]
6: 2886010 Solidere
[pem_pos=0, final_scores=0.6854134,
final_local_scores=0.67219955, pem=0.0, similarity=24.149488]
7: 9848049 Beirut_Stock_Exchange
[pem_pos=0, final_scores=0.32863742,
final_local_scores=0.41797388, pem=0.0, similarity=16.562887]
8: 27417693 Eternit
[pem_pos=0, final_scores=0.44435346,
final_local_scores=0.6290216, pem=0.0, similarity=22.860971]
9: 37428 Beirut
[pem_pos=0, final_scores=0.476587,
final_local_scores=0.4750617, pem=-0.046614014,
similarity=18.519997]

Sample: b'209_Zenith'

[1 1 Zenith 1] lands \$ 1 billion contract , plans \$ 100 million plant . Susan Nadeau [2 2 CHICAGO 2] 1996-08-22 A

consortium of telephone companies and The Walt Disney Co . said Thursday it had signed a \$ 1 billion contract with [3 3 Zenith 3] to make digital television set-top boxes for its home entertainment service . The announcement of the contract for 3 million set-top boxes gave new hope to [4 4 Zenith 4] , which has struggled with years of losses . " This really indicates we 're back in the business , " William Luehrs , president of the Glenview , Ill.-based company 's Networks Services Division , said in a telephone interview . " Anytime somebody gets the opportunity to enter the next era with such a big bang has got to be seen as a strong message to the industry . " Following the announcement , [5 5 Zenith 5] 's stock soared \$ 5.50 to \$ 16.875 on the [6 6 New York Stock Exchange 6] . The consortium , called Americast , said the contract was part of its strategy to develop and market the next generation in home entertainment . In addition to [7 7 Disney 7] , Americast 's partners are phone companies Ameritech Corp. , 1 BellSouth Corp. , GTE Corp. and SBC Communications . Americast said 2 Southern New England Telecommunications Corp. has signed a letter of intent to join the group , which plans to provide a home entertainment service similar to cable television . [8 8 Zenith 8] also said it planned to build a new \$ 100 million plant in [9 9 Woodridge 9] , [10 10 Ill 10] . , to make picture tubes for 32- and 35-inch screen TV sets . The company currently buys the tubes from competitors . The new plant , which is dependent on obtaining financing , will create about 280 new jobs , [11 11 Zenith 11] said . The contract calls for production of the set-top boxes over five years . Luehrs said manufacturing will begin and revenue will start to roll in during the first half of next year . The boxes will be made on a build-to-order basis . [12 12 Zenith 12] will convert its [13 13 Chihuahua 13] , [14 14 Mexico 14] , analogue set-top box plant to manufacture the digital boxes . Luehrs declined to say when the operation was expected to be profitable . Americast will provide the boxes to subscribers as part of the service . Its service is being introduced in selected markets across the [15 15 United States 15] . Luehrs said digital technology in set-top boxes is only the beginning and said the technology will eventually show up in retail consumer electronics . " We 'll build a lot of these devices into television sets , for digital television , " he said . " Although that is not where this particular contract is headed , the fact that there is a strong [16 16 Zenith 16] presence will pay us dividends in the future . " [17 17 Zenith 17] has been plagued by generally soft conditions in the colour television industry , reporting full-year losses since 1989 . Last month , it reported a second-quarter loss of \$ 33.2 million , or 51 cents a share , vs. a loss of \$ 45.3 million , or 97 cents a share , a year earlier . Last November , South Korea-based LG Electronics Inc. bought a majority stake in [18 18 Zenith 18] . Robert Gutenstein , an analyst for Kalf , Voorhis & Co . ,

said the contract was " not unique , but it 's big . " "
Digital is coming , it 's economic and the question is what
will make the consumer happy and at what price . "

1: 899786 Zenith_Electronics
2: 6886 Chicago
3: 899786 Zenith_Electronics
4: 899786 Zenith_Electronics
5: 899786 Zenith_Electronics
6: 21560 New_York_Stock_Exchange
7: 37398 The_Walt_Disney_Company
8: 899786 Zenith_Electronics
9: 111071 Woodridge,_Illinois
10: 14849 Illinois
11: 899786 Zenith_Electronics
12: 899786 Zenith_Electronics
13: 217560 Chihuahua,_Chihuahua
14: 3966054 Mexico
15: 3434750 United_States
16: 899786 Zenith_Electronics
17: 899786 Zenith_Electronics
18: 899786 Zenith_Electronics

1: 335382 BellSouth
[winner: 7485 Corporation, pem_pos=0,
final_scores=-1.1553842, final_local_scores=-1.2622647,
pem=-0.005830923, similarity=-33.547073]
RECALL MISS: 335382 BellSouth]
[winner: 335382 BellSouth, pem_pos=0,
final_scores=-1.3158722, final_local_scores=-1.4383483,
pem=0.0, similarity=-38.83347]
gt: 335382 BellSouth, pem_pos=0,
final_scores=-1.3158722, final_local_scores=-1.4383483,
pem=0.0, similarity=-38.83347]
2: 340894 Southern_New_England_Telecommunications
[winner: 340894 Southern_New_England_Telecommunications,
pem_pos=0, final_scores=-0.98807335,
final_local_scores=-0.9919644, pem=0.0,
similarity=-25.512484]
gt: 340894 Southern_New_England_Telecommunications,
pem_pos=0, final_scores=-0.98807335,
final_local_scores=-0.9919644, pem=0.0,
similarity=-25.512484]
[winner: 7485 Corporation, pem_pos=0,
final_scores=-1.1044853, final_local_scores=-1.2066653,
pem=-0.005830923, similarity=-31.887877]
RECALL MISS: 340894
Southern_New_England_Telecommunications]
1: 899786 Zenith_Electronics
[pem_pos=2, final_scores=0.5656702,
final_local_scores=0.52584255, pem=-1.6902313,
similarity=28.973644]
2: 6886 Chicago
[pem_pos=0, final_scores=0.5884083,

```
final_local_scores=0.62496096, pem=-0.20658566,
similarity=23.863241]
3: 899786 Zenith_Electronics
    [pem_pos=2, final_scores=0.59597033,
final_local_scores=0.5589409, pem=-1.6902313,
similarity=29.961363]
4: 899786 Zenith_Electronics
    [pem_pos=2, final_scores=0.7612174,
final_local_scores=0.7394483, pem=-1.6902313,
similarity=35.348064]
5: 899786 Zenith_Electronics
    [pem_pos=2, final_scores=0.6906839,
final_local_scores=0.6624011, pem=-1.6902313,
similarity=33.048824]
6: 21560 New_York_Stock_Exchange
    [pem_pos=0, final_scores=0.42762572,
final_local_scores=0.4699366, pem=0.0, similarity=18.113558]
7: 37398 The_Walt_Disney_Company
    [pem_pos=1, final_scores=0.48507482,
final_local_scores=0.53944063, pem=-0.24767768,
similarity=21.534607]
8: 899786 Zenith_Electronics
    [pem_pos=2, final_scores=0.59665763,
final_local_scores=0.55969167, pem=-1.6902313,
similarity=29.983768]
9: 111071 Woodridge,_Illinois
    [pem_pos=0, final_scores=0.4484015,
final_local_scores=0.5680243, pem=-0.37740436,
similarity=23.093073]
10: 14849 Illinois
    [pem_pos=2, final_scores=0.42987972,
final_local_scores=0.47328454, pem=-2.0570593,
similarity=29.400076]
11: 899786 Zenith_Electronics
    [pem_pos=2, final_scores=0.58747655,
final_local_scores=0.5496627, pem=-1.6902313,
similarity=29.684484]
12: 899786 Zenith_Electronics
    [pem_pos=2, final_scores=0.5862043,
final_local_scores=0.54827297, pem=-1.6902313,
similarity=29.643013]
13: 217560 Chihuahua,_Chihuahua
    [pem_pos=2, final_scores=0.21935083,
final_local_scores=0.30108565, pem=-1.4471421,
similarity=20.944494]
14: 3966054 Mexico
    [pem_pos=0, final_scores=0.49529624,
final_local_scores=0.6054874, pem=-0.18258883,
similarity=23.151611]
15: 3434750 United_States
    [pem_pos=0, final_scores=0.40699744,
final_local_scores=0.45096123, pem=-0.073457584,
```

```
similarity=17.94677]
16: 899786 Zenith_Electronics
    [pem_pos=2, final_scores=0.5175641,
final_local_scores=0.47329402, pem=-1.6902313,
similarity=27.40549]
17: 899786 Zenith_Electronics
    [pem_pos=2, final_scores=0.65255684,
final_local_scores=0.62075305, pem=-1.6902313,
similarity=31.805964]
18: 899786 Zenith_Electronics
    [pem_pos=2, final_scores=0.6440594,
final_local_scores=0.61147094, pem=-1.6902313,
similarity=31.528965]

Sample: b'210_Natural'
[1 1 Natural Law Party 1] says it can meditate problems
away . [2 2 WASHINGTON 2] 1996-08-22 From the people who
brought you hundreds of " yogic fliers " who claimed to defy
nature by levitating comes The Natural Law Party , a minor
political party that nominated a presidential candidate on
Thursday . At a hotel convention here , the party associated
with the [3 3 Transcendental Meditation 3] ( TM ) movement
named physicist [4 4 John Hagelin 4] as its presidential
nominee for the Nov. 5 election . The party is running on a
platform claiming it can ward off problems before they occur
through techniques such as mass meditation that would reduce
stress , crime , terrorism and even wars . " Social stress
can be reduced and problems such as crime and violence will
automatically decrease , " said a party paper . Many party
members are practitioners of TM , which involves meditating
to a repeated word or phrase , called a mantra . Some
advanced TM followers contend they can actually mediate to
such a point that they fly . But in a demonstration of "
yogic flying " several years ago , critics said the people
were merely bouncing off the ground from a sitting position .
1: 32012 Natural_Law_Party_(United_States)
2: 108956 Washington,_D.C.
3: 42734 Transcendental_Meditation
4: 16077 John_Hagelin
1: 32012 Natural_Law_Party_(United_States)
    [pem_pos=1, final_scores=0.29399842,
final_local_scores=0.30600095, pem=-0.38415533,
similarity=15.310487]
2: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.54129773,
final_local_scores=0.62212735, pem=-1.7496481,
similarity=32.170094]
3: 42734 Transcendental_Meditation
    [pem_pos=0, final_scores=0.44236082,
final_local_scores=0.48505014, pem=0.0, similarity=18.564577]
4: 16077 John_Hagelin
    [pem_pos=0, final_scores=0.6388684,
```

final_local_scores=0.66440856, pem=0.0, similarity=23.91699]

Sample: b'211_Huge'

Huge [1 1 Windows 95 1] sales fail to meet expectations . Martin Wolk [2 2 SEATTLE 2] 1996-08-22 A year after its massively publicized introduction , [3 3 Microsoft Corp. 3] 's [4 4 Windows 95 4] computer operating system has fallen short of the most optimistic expectations for the software giant and the industry . Even though more than 40 million copies of [5 5 Windows 95 5] have been sold , making it the fastest-selling new software ever , it would have been impossible for any product to live up to the unprecedented hype of the Aug. 24 , 1995 launch , when stores around the world opened at midnight to greet long lines of customers . The [6 6 Redmond 6] , Wash.-based company spent tens of millions of dollars promoting the product with stunts that included buying the entire print run of the [7 7 Times of London 7] and lighting [8 8 New York 8] 's [9 9 Empire State 9] building in a [10 10 Windows 10] color scheme . But the product , delivered eight months late , has fallen short of its sales potential in part because [11 11 Microsoft 11] delivered a mixed message to business customers , analysts said . " It did n't do as well as it could have , " said [12 1 2 Rob Enderle 12] , an analyst with Giga Information Group . Scores of software and hardware companies that had hoped for a big boost in sales were disappointed when only a brief spike materialized . " People who were expecting major coat-tails were somewhat disappointed , " said Scott Winkler , an analyst with [13 13 Gartner Group 13] . " It 's not as though it has n't had an impact , " he said . " It just has n't had the huge earth-shattering impact some people were looking for . " Symantec Corp. , which had been among the most bullish of software companies at the time of the [14 14 Windows 95 14] launch , ended up posting disappointing financial results when retail sales of the operating system fell short of its projections . Touchstone Software Corp. had to pay \$ 1.3 million in cash and stock to settle a shareholders lawsuit brought after the company 's sales failed to meet expectations tied to the [15 15 Windows 95 15] launch . Many software developers apparently saw their crucial holiday season sales suffer last year because store shelves were jammed with blue-and-white boxes of [16 16 Windows 95 16] , resulting in a shortage of space for seasonal products , said Ann Stephens , president of PC Data Inc . To be sure , sales of [17 17 Windows 95 17] and the accompanying Office 95 upgrade drove [18 18 Microsoft 18] sales up 46 percent last year to a record \$ 8.67 billion and cemented the company 's status as the industry 's dominant company . [19 19 Microsoft 19] executives say they are thrilled with the sales figures , and industry analysts estimate that by sometime next year , the installed base of [20 20 Windows 95 20] will surpass that of the older version of [21 21 Windows 21] , now used on

about 100 million computers worldwide . But [22 22 Enderle 22] said the figure could have been even higher if [23 23 Microsoft 23] had done a better job of handling the huge demand for technical support from customers who were frustrated trying to install the system . He and other analysts said corporate [24 24 America 24] adopted a go-slow approach because [25 25 Microsoft 25] already was promoting the new version of its high-end [26 26 Windows NT 26] operating system , expected to be available in stores in the next several weeks . " [27 27 Microsoft 27] sent a lot of signals that NT was going to be the answer , " Winkler said . " Many people began to believe that [28 28 Windows 95 28] was being downplayed . " But now that [29 29 Windows NT 4.0 29] has been launched , Winkler and others believe only a relatively small proportion of corporate users will elect to pay the added software and hardware costs needed to use it instead of [30 30 Windows 95 30] . " [31 31 Windows 95 31] is going to do great , " he said . " The mistake people made was in thinking it was going to be a fast , sweeping change rather than a slow , building change . "

1: 34064 Windows_95
2: 11388236 Seattle
3: 19001 Microsoft
4: 34064 Windows_95
5: 34064 Windows_95
6: 26558 Redmond,_Washington
7: 39127 The_Times
8: 645042 New_York_City
9: 9736 Empire_State_Building
10: 18890 Microsoft_Windows
11: 19001 Microsoft
12: 4920586 Rob_Enderle
13: 250363 Gartner
14: 34064 Windows_95
15: 34064 Windows_95
16: 34064 Windows_95
17: 34064 Windows_95
18: 19001 Microsoft
19: 19001 Microsoft
20: 34064 Windows_95
21: 18890 Microsoft_Windows
22: 4920586 Rob_Enderle
23: 19001 Microsoft
24: 3434750 United_States
25: 19001 Microsoft
26: 21291483 Windows_NT
27: 19001 Microsoft
28: 34064 Windows_95
29: 795680 Windows_NT_4.0
30: 34064 Windows_95
31: 34064 Windows_95
1: 34064 Windows_95

```
[pem_pos=0, final_scores=0.42496175,
final_local_scores=0.3696226, pem=0.0, similarity=15.1199875]
2: 11388236 Seattle
    [pem_pos=0, final_scores=0.43311512,
final_local_scores=0.5033876, pem=-0.1054736,
similarity=19.685385]
3: 19001 Microsoft
    [pem_pos=0, final_scores=0.44018376,
final_local_scores=0.38554186, pem=-0.006739087,
similarity=15.6317005]
4: 34064 Windows_95
    [pem_pos=0, final_scores=0.45525628,
final_local_scores=0.40271473, pem=0.0, similarity=16.107523]
5: 34064 Windows_95
    [pem_pos=0, final_scores=0.44207513,
final_local_scores=0.38831633, pem=0.0, similarity=15.677847]
6: 26558 Redmond,_Washington
    [pem_pos=0, final_scores=0.43604076,
final_local_scores=0.48569572, pem=-0.31218576,
similarity=20.281557]
7: 39127 The_Times
    [pem_pos=0, final_scores=0.44751996,
final_local_scores=0.5506456, pem=0.0, similarity=20.522076]
8: 645042 New_York_City
    [pem_pos=1, final_scores=0.3766513,
final_local_scores=0.44919473, pem=-1.7906687,
similarity=27.232513]
9: 9736 Empire_State_Building
    [pem_pos=0, final_scores=0.38204277,
final_local_scores=0.48942685, pem=0.0, similarity=18.695187]
10: 18890 Microsoft_Windows
    [pem_pos=0, final_scores=0.42649525,
final_local_scores=0.37990582, pem=-0.047305938,
similarity=15.684116]
11: 19001 Microsoft
    [pem_pos=0, final_scores=0.5326373,
final_local_scores=0.48653328, pem=-0.006739087,
similarity=18.645485]
12: 4920586 Rob_Enderle
    [pem_pos=0, final_scores=0.49830472,
final_local_scores=0.6177607, pem=0.0, similarity=22.524925]
13: 250363 Gartner
    [pem_pos=0, final_scores=0.51214355,
final_local_scores=0.54738563, pem=0.0, similarity=20.424791]
14: 34064 Windows_95
    [pem_pos=0, final_scores=0.43653613,
final_local_scores=0.38226587, pem=0.0, similarity=15.49729]
15: 34064 Windows_95
    [pem_pos=0, final_scores=0.44944936,
final_local_scores=0.3963716, pem=0.0, similarity=15.918232]
16: 34064 Windows_95
    [pem_pos=0, final_scores=0.4090219,
```

```
final_local_scores=0.3522107, pem=0.0, similarity=14.600383]
17: 34064 Windows_95
    [pem_pos=0, final_scores=0.461577,
final_local_scores=0.40961915, pem=0.0, similarity=16.313566]
18: 19001 Microsoft
    [pem_pos=0, final_scores=0.45244986,
final_local_scores=0.39894074, pem=-0.006739087,
similarity=16.03155]
19: 19001 Microsoft
    [pem_pos=0, final_scores=0.50068396,
final_local_scores=0.4516291, pem=-0.006739087,
similarity=17.603874]
20: 34064 Windows_95
    [pem_pos=0, final_scores=0.4478491,
final_local_scores=0.39462352, pem=0.0, similarity=15.866066]
21: 18890 Microsoft_Windows
    [pem_pos=0, final_scores=0.30915374,
final_local_scores=0.25172806, pem=-0.047305938,
similarity=11.859037]
22: 4920586 Rob_Enderle
    [pem_pos=0, final_scores=0.694845,
final_local_scores=0.8440201, pem=0.0, similarity=29.276957]
23: 19001 Microsoft
    [pem_pos=0, final_scores=0.50294954,
final_local_scores=0.4541039, pem=-0.006739087,
similarity=17.677727]
24: 3434750 United_States
    [pem_pos=0, final_scores=0.4823205,
final_local_scores=0.5654814, pem=-0.78483206,
similarity=25.232845]
25: 19001 Microsoft
    [pem_pos=0, final_scores=0.4761141,
final_local_scores=0.42479032, pem=-0.006739087,
similarity=16.802952]
26: 21291483 Windows_NT
    [pem_pos=0, final_scores=0.27552748,
final_local_scores=0.23672357, pem=0.0, similarity=11.1540165]
27: 19001 Microsoft
    [pem_pos=0, final_scores=0.4663934,
final_local_scores=0.41417193, pem=-0.006739087,
similarity=16.486078]
28: 34064 Windows_95
    [pem_pos=0, final_scores=0.43585157,
final_local_scores=0.38151807, pem=0.0, similarity=15.474974]
29: 795680 Windows_NT_4.0
    [pem_pos=0, final_scores=0.31090575,
final_local_scores=0.29239026, pem=0.0, similarity=12.815222]
30: 34064 Windows_95
    [pem_pos=0, final_scores=0.40848005,
final_local_scores=0.35161883, pem=0.0, similarity=14.58272]
31: 34064 Windows_95
    [pem_pos=0, final_scores=0.45411086,
```

final_local_scores=0.40146357, pem=0.0, similarity=16.070187]

Sample: b'212_U.S.'

[1 1 U.S. 1] says still committed to [2 2 Cuba 2] migration pacts . [3 3 WASHINGTON 3] 1996-08-22 The [4 4 United States 4] said on Thursday it remained committed to migration accords with [5 5 Cuba 5] and would continue to repatriate intercepted [6 6 Cuban 6] migrants who attempted to enter [7 7 U.S. 7] territory illegally . A [8 8 State Department 8] statement appeared in part a response to [9 9 Cuban 9] complaints that 1 Washington was jeopardising the accords by failing to return some of the Cubans involved in recent illegal migration incidents . " The [10 10 United States 10] reiterates its full commitment to the implementation " of the accords signed by the two countries in 1994 and 1995 , said the statement by spokesman Glyn Davies . " The [11 11 United States 11] will continue to return [12 12 Cuban 12] migrants intercepted at sea who seek to enter the [13 13 United States 13] or the Guantanamo Naval Base illegally , " it said . 2 Washington would also take " prompt and effective law enforcement action " against alien smuggling and hijackings from [14 14 Cuba 14] , it added . Davies told reporters the statement would be distributed in the [15 15 Cuban 15] exile community in [16 16 Miami 16] " to remind everyone of the importance of abiding by the accords and avoiding dangerous attempts to cross the straits " from [17 17 Cuba 17] to Florida . Havana 's complaints centred on an incident in which a boatload of emigrants capsized in the [18 18 Florida Straits 18] last week and two recent aircraft hijackings from [19 19 Cuba 19] . Sixteen of those picked up from the boat were returned to [20 20 Cuba 20] but eight were taken to the [21 21 United States 21] and three to Guantanamo Bay , a [22 22 U.S. 22] base on [23 23 Cuba 23] , until they emigrated to another nation . The most recent hijacking , last Friday , involved three hijackers and the pilot of a small aircraft . Davies said on Tuesday the pilot could soon return to [24 24 Cuba 24] but [25 25 U.S. 25] authorities planned to try the hijackers . In an incident on July 7 , a [26 26 Cuban 26] [27 27 interior ministry 27] official hijacked a commercial plane and sought asylum at Guantanamo Bay . Davies said he knew of no plans to return the man to [28 28 Cuba 28] .

1: 3434750 United_States

2: 5042481 Cuba

3: 108956 Washington,_D.C.

4: 3434750 United_States

5: 5042481 Cuba

6: 5042481 Cuba

7: 3434750 United_States

8: 31975 United_States_Department_of_State

9: 5042481 Cuba

10: 3434750 United_States

11: 3434750 United_States

```
12: 5042481 Cuba
13: 3434750 United_States
14: 5042481 Cuba
15: 5042481 Cuba
16: 53846 Miami
17: 5042481 Cuba
18: 461620 Straits_of_Florida
19: 5042481 Cuba
20: 5042481 Cuba
21: 3434750 United_States
22: 3434750 United_States
23: 5042481 Cuba
24: 5042481 Cuba
25: 3434750 United_States
26: 5042481 Cuba
27: 354620 Interior_ministry
28: 5042481 Cuba
1: 3434750 United_States
    [winner: 108956 Washington,_D.C., pem_pos=0,
final_scores=-0.17361055, final_local_scores=-0.243983,
pem=-1.7496481, similarity=6.3236313|
     RECALL MISS: 3434750 United_States]
2: 3434750 United_States
    [winner: 108956 Washington,_D.C., pem_pos=0,
final_scores=-0.089112595, final_local_scores=-0.15168181,
pem=-1.7496481, similarity=9.078083|
     RECALL MISS: 3434750 United_States]
1: 3434750 United_States
    [pem_pos=0, final_scores=0.51236314,
final_local_scores=0.46777648, pem=-0.14075686,
similarity=18.814552]
2: 5042481 Cuba
    [pem_pos=0, final_scores=0.55739623,
final_local_scores=0.5000451, pem=-0.23907764,
similarity=20.312199]
3: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.5951219,
final_local_scores=0.6013201, pem=-1.7496481,
similarity=31.54916]
4: 3434750 United_States
    [pem_pos=0, final_scores=0.52242655,
final_local_scores=0.47876924, pem=-0.073457584,
similarity=18.776617]
5: 5042481 Cuba
    [pem_pos=0, final_scores=0.52583045,
final_local_scores=0.4655643, pem=-0.23907764,
similarity=19.28322]
6: 5042481 Cuba
    [pem_pos=0, final_scores=0.50427824,
final_local_scores=0.4420218, pem=-0.37537703,
similarity=19.321884]
7: 3434750 United_States
```

```
[pem_pos=0, final_scores=0.5272632,
final_local_scores=0.4840526, pem=-0.14075686,
similarity=19.300266]
8: 31975 United_States_Department_of_State
    [pem_pos=0, final_scores=0.34920216,
final_local_scores=0.35739735, pem=0.0, similarity=14.755162]
9: 5042481 Cuba
    [pem_pos=0, final_scores=0.57009417,
final_local_scores=0.5139157, pem=-0.37537703,
similarity=21.467342]
10: 3434750 United_States
    [pem_pos=0, final_scores=0.51912993,
final_local_scores=0.47516817, pem=-0.073457584,
similarity=18.669153]
11: 3434750 United_States
    [pem_pos=0, final_scores=0.52439076,
final_local_scores=0.48091483, pem=-0.073457584,
similarity=18.840645]
12: 5042481 Cuba
    [pem_pos=0, final_scores=0.49900478,
final_local_scores=0.43626136, pem=-0.37537703,
similarity=19.14998]
13: 3434750 United_States
    [pem_pos=0, final_scores=0.5049628,
final_local_scores=0.45969272, pem=-0.073457584,
similarity=18.207335]
14: 5042481 Cuba
    [pem_pos=0, final_scores=0.5630371,
final_local_scores=0.5062069, pem=-0.23907764,
similarity=20.49608]
15: 5042481 Cuba
    [pem_pos=0, final_scores=0.51313525,
final_local_scores=0.4516967, pem=-0.37537703,
similarity=19.610601]
16: 53846 Miami
    [pem_pos=0, final_scores=0.46723545,
final_local_scores=0.46817535, pem=-0.35396606,
similarity=19.985922]
17: 5042481 Cuba
    [pem_pos=0, final_scores=0.5311853,
final_local_scores=0.47141367, pem=-0.23907764,
similarity=19.457777]
18: 461620 Straits_of_Florida
    [pem_pos=0, final_scores=0.39245903,
final_local_scores=0.47046936, pem=0.0, similarity=18.129457]
19: 5042481 Cuba
    [pem_pos=0, final_scores=0.58386743,
final_local_scores=0.5289609, pem=-0.23907764,
similarity=21.175102]
20: 5042481 Cuba
    [pem_pos=0, final_scores=0.50473684,
final_local_scores=0.4425227, pem=-0.23907764,
```

```
similarity=18.595613]
21: 3434750 United_States
    [pem_pos=0, final_scores=0.5473715,
final_local_scores=0.5060178, pem=-0.073457584,
similarity=19.589767]
22: 3434750 United_States
    [pem_pos=0, final_scores=0.52808285,
final_local_scores=0.48494786, pem=-0.14075686,
similarity=19.326982]
23: 5042481 Cuba
    [pem_pos=0, final_scores=0.55363536,
final_local_scores=0.49593693, pem=-0.23907764,
similarity=20.189602]
24: 5042481 Cuba
    [pem_pos=0, final_scores=0.53066564,
final_local_scores=0.470846, pem=-0.23907764,
similarity=19.440838]
25: 3434750 United_States
    [pem_pos=0, final_scores=0.49549693,
final_local_scores=0.44935268, pem=-0.14075686,
similarity=18.264751]
26: 5042481 Cuba
    [pem_pos=0, final_scores=0.55642146,
final_local_scores=0.49898028, pem=-0.37537703,
similarity=21.021639]
27: 354620 Interior_ministry
    [pem_pos=0, final_scores=0.54152405,
final_local_scores=0.6297112, pem=-0.57122487,
similarity=25.98796]
28: 5042481 Cuba
    [pem_pos=0, final_scores=0.58717495,
final_local_scores=0.5325739, pem=-0.23907764,
similarity=21.282923]
```

Sample: b'213_Wis'

Wis . says is first state to apply for new welfare . [1 1 CH ICAGO 1] 1996-08-22 In keeping with its pioneering image in the area of welfare , [2 2 Wisconsin 2] was the first state to submit an administrative plan under the nation 's new welfare law , Gov . [3 3 Tommy Thompson 3] said Thursday . According to a new release from the governor , [4 4 Wisconsin 4] submitted a plan to the U.S. Department of Health and Human Services for administration of the new block grant system for welfare just minutes after President [5 5 Bill Clinton 5] signed the measure into law Thursday . " As the nation 's leader in welfare reform , [6 6 Wisconsin 6] is far ahead of the curve and ready to go under this new system , " [7 7 Thompson 7] said . Still , the governor said the new law does not go as far as the state 's own welfare reform program , dubbed W-2 . He said that despite the new law , [8 8 Wisconsin 8] will still require federal waivers allowing the working poor to acquire health care coverage from the state ,

a 60-day residency requirement for participation in the welfare program , and child support collections to go directly to custodial parents . The nation 's new welfare reform law limits eligibility , gives states more power and ends direct federal aid for poor children . -- Karen Pierog , 312-408-8647

1: 6886 Chicago
2: 33127 Wisconsin
3: 217709 Tommy_Thompson
4: 33127 Wisconsin
5: 3356 Bill_Clinton
6: 33127 Wisconsin
7: 217709 Tommy_Thompson
8: 33127 Wisconsin
1: 6886 Chicago
[pem_pos=0, final_scores=0.6347229,
final_local_scores=0.64533454, pem=-0.20658566,
similarity=24.471228]
2: 33127 Wisconsin
[pem_pos=0, final_scores=0.5665843,
final_local_scores=0.5226105, pem=-0.21268336,
similarity=20.842056]
3: 217709 Tommy_Thompson
[pem_pos=0, final_scores=0.40545094,
final_local_scores=0.41341674, pem=-0.041727614,
similarity=16.653814]
4: 33127 Wisconsin
[pem_pos=0, final_scores=0.546313,
final_local_scores=0.5004672, pem=-0.21268336,
similarity=20.181257]
5: 3356 Bill_Clinton
[pem_pos=0, final_scores=0.43277496,
final_local_scores=0.46434307, pem=0.0, similarity=17.946636]
6: 33127 Wisconsin
[pem_pos=0, final_scores=0.6201735,
final_local_scores=0.58114856, pem=-0.21268336,
similarity=22.58895]
7: 217709 Tommy_Thompson
[pem_pos=0, final_scores=0.6528053,
final_local_scores=0.68361384, pem=-0.041727614,
similarity=24.717033]
8: 33127 Wisconsin
[pem_pos=0, final_scores=0.6217647,
final_local_scores=0.58288676, pem=-0.21268336,
similarity=22.640821]

Sample: b'214_Snoozing'
Snoozing [1 1 Vietnamese 1] man takes slow train to [2 2 Alaska 2] . [3 3 ANCHORAGE 3] , [4 4 Alaska 4] 1996-08-22 A [5 5 Vietnamese 5] man who tried to take a snooze in a railway boxcar in [6 6 Canada 6] found himself locked in and bound for [7 7 Alaska 7] with no food or water . Officials in the

port of [8 8 Whittier 8] said on Thursday that they found Tuan Quac Phan , 29 , dehydrated , famished and terrified after sailing to [9 9 Alaska 9] from [10 10 Canada 10] in the boxcar loaded on a barge , a trip that takes about five days . Sgt . Dan Jewell of the [11 11 Whittier 11] , [12 12 Alaska 12] police department described Phan as " extremely cooperative " . " Seeing me in my uniform , he kept saying , Jail better . Jail better . ' " Phan 's accidental journey started last week in [13 13 Prince Rupert 13] , [14 14 British Columbia 14] , where he was searching for a fishing job , Jewell said . " He had climbed up in this boxcar to get out of the weather and to get some sleep , " Jewell said . " The next thing you know , the boxcar is coupled up and loaded up to a barge and headed north . " Police found Phan late on Monday when the boxcar , which was transporting lumber , was opened at [15 15 Whittier 15] , a port in western [16 16 Prince William Sound 16] . Officials fed Phan some soup , gave him medical care , kept him overnight and then fed him a large breakfast .

1: 202354 Vietnam
2: 624 Alaska
3: 7004698 Anchorage,_Alaska
4: 624 Alaska
5: 202354 Vietnam
6: 5042916 Canada
7: 624 Alaska
8: 105698 Whittier,_Alaska
9: 624 Alaska
10: 5042916 Canada
11: 105698 Whittier,_Alaska
12: 624 Alaska
13: 385576 Prince_Rupert,_British_Columbia
14: 3392 British_Columbia
15: 105698 Whittier,_Alaska
16: 197927 Prince_William_Sound
1: 202354 Vietnam
[pem_pos=1, final_scores=0.622206,
final_local_scores=0.67076945, pem=-1.579876,
similarity=32.69842]
2: 624 Alaska
[pem_pos=0, final_scores=0.49434817,
final_local_scores=0.43818378, pem=-0.18647984,
similarity=18.180098]
3: 7004698 Anchorage,_Alaska
[pem_pos=0, final_scores=0.50573176,
final_local_scores=0.47721905, pem=-0.046072915,
similarity=18.581434]
4: 624 Alaska
[pem_pos=0, final_scores=0.5323942,
final_local_scores=0.47974324, pem=-0.18647984,
similarity=19.420315]
5: 202354 Vietnam

```
[pem_pos=1, final_scores=0.5968534,
final_local_scores=0.6430756, pem=-1.579876,
similarity=31.87198]
6: 5042916 Canada
    [pem_pos=0, final_scores=0.7678358,
final_local_scores=0.7874195, pem=-0.11662793,
similarity=28.222122]
7: 624 Alaska
    [pem_pos=0, final_scores=0.62003297,
final_local_scores=0.57547534, pem=-0.18647984,
similarity=22.277151]
8: 105698 Whittier,_Alaska
    [pem_pos=1, final_scores=0.41336018,
final_local_scores=0.41503662, pem=-1.2735102,
similarity=23.40078]
9: 624 Alaska
    [pem_pos=0, final_scores=0.67752296,
final_local_scores=0.63827443, pem=-0.18647984,
similarity=24.151201]
10: 5042916 Canada
    [pem_pos=0, final_scores=0.74847716,
final_local_scores=0.75639313, pem=-0.11662793,
similarity=27.296234]
11: 105698 Whittier,_Alaska
    [pem_pos=1, final_scores=0.49695665,
final_local_scores=0.5063531, pem=-1.2735102,
similarity=26.125845]
12: 624 Alaska
    [pem_pos=0, final_scores=0.7221031,
final_local_scores=0.6869715, pem=-0.18647984,
similarity=25.604418]
13: 385576 Prince_Rupert,_British_Columbia
    [pem_pos=0, final_scores=0.39872068,
final_local_scores=0.3940199, pem=-0.6844881,
similarity=19.570406]
14: 3392 British_Columbia
    [pem_pos=0, final_scores=0.5020625,
final_local_scores=0.46979785, pem=-0.028628228,
similarity=18.265102]
15: 105698 Whittier,_Alaska
    [pem_pos=1, final_scores=0.44522566,
final_local_scores=0.44984478, pem=-1.2735102,
similarity=24.439526]
16: 197927 Prince_William_Sound
    [pem_pos=0, final_scores=0.3770662,
final_local_scores=0.39838552, pem=0.0, similarity=15.978332]
```

Sample: b'215_State'

State , federal agents probe [1 1 Arkansas 1] church fires .
Steve Barnes [2 2 LITTLE ROCK 2] , Ark . 1996-08-22 State and
federal agents on Thursday sifted through the rubble of two
predominantly black [3 3 Arkansas 3] churches that burned

within minutes of one another late Tuesday and early Wednesday . Both churches were in the [4 4 Mississippi 4] delta region of [5 5 Arkansas 5] , about 90 miles (145 kms) southeast of [6 6 Little Rock 6] , and were located within three miles of one another . " We 're investigating with the idea that both fires may be arson , but that has n't been conclusively established , " said Wayne Jordan , a spokesman for the [7 7 Arkansas State Police 7] . Agents of the [8 8 F. B.I. 8] and the Bureau of Alcohol , Tobacco and Firearms were also at the scene , Jordan said . Mount Zion Missionary Baptist Church and St. Matthews Missionary Baptist Church were both frame structures , each near Turner , [9 9 Arkansas 9] , a small community surrounded by cotton and soybean fields . " This is rural [10 10 Arkansas 10] . I 'm surprised anyone could even find us out here , " said Fannie Johnson , a member of St. Matthew 's , who said she believed arson was to blame . Others connected with the two churches said they shared that suspicion , although all said they knew of no motive and no racial tension in the area . " It 's sad someone would have that kind of spite in their heart , " said Rev. Jerome Turner , pastor of Mount Zion Missionary Baptist Church . [11 11 Arkansas 11] has been spared the loss of predominantly black churches to arson , a wave that has claimed an estimated 30 houses of worship across the south in the past several months . A black church near [12 12 Camden 12] , [13 13 Arkansas 13] , about 100 miles (161 kms) south of [14 14 Little Rock 14] , burned in July , but federal agents have not determined the cause .

1: 1930 Arkansas
2: 18520327 Little_Rock,_Arkansas
3: 1930 Arkansas
4: 19579 Mississippi_River
5: 1930 Arkansas
6: 18520327 Little_Rock,_Arkansas
7: 7945316 Arkansas_State_Police
8: 11127 Federal_Bureau_of_Investigation
9: 1930 Arkansas
10: 1930 Arkansas
11: 1930 Arkansas
12: 107167 Camden,_Arkansas
13: 1930 Arkansas
14: 18520327 Little_Rock,_Arkansas
1: 1930 Arkansas
[pem_pos=0, final_scores=0.7063002,
final_local_scores=0.65453595, pem=-0.32007957,
similarity=25.363014]
2: 18520327 Little_Rock,_Arkansas
[pem_pos=0, final_scores=0.56734216,
final_local_scores=0.5318367, pem=-0.1562997,
similarity=20.810762]
3: 1930 Arkansas
[pem_pos=0, final_scores=0.6129622,

```
final_local_scores=0.5525784, pem=-0.32007957,
similarity=22.320396]
4: 19579 Mississippi_River
    [pem_pos=1, final_scores=0.3494535,
final_local_scores=0.35507834, pem=-2.3293974,
similarity=27.353584]
5: 1930 Arkansas
    [pem_pos=0, final_scores=0.55449164,
final_local_scores=0.48870814, pem=-0.32007957,
similarity=20.414381]
6: 18520327 Little_Rock,_Arkansas
    [pem_pos=0, final_scores=0.42455298,
final_local_scores=0.37586123, pem=-0.1562997,
similarity=16.156143]
7: 7945316 Arkansas_State_Police
    [pem_pos=0, final_scores=0.4417007,
final_local_scores=0.532042, pem=0.0, similarity=19.966908]
8: 11127 Federal_Bureau_of_Investigation
    [pem_pos=0, final_scores=0.4466532,
final_local_scores=0.5184764, pem=0.0, similarity=19.562084]
9: 1930 Arkansas
    [pem_pos=0, final_scores=0.582639,
final_local_scores=0.51945484, pem=-0.32007957,
similarity=21.331924]
10: 1930 Arkansas
    [pem_pos=0, final_scores=0.6546107,
final_local_scores=0.59807307, pem=-0.32007957,
similarity=23.67805]
11: 1930 Arkansas
    [pem_pos=0, final_scores=0.5667023,
final_local_scores=0.5020464, pem=-0.32007957,
similarity=20.812422]
12: 107167 Camden,_Arkansas
    [pem_pos=5, final_scores=0.34612954,
final_local_scores=0.31981653, pem=-2.991406,
similarity=29.901407]
13: 1930 Arkansas
    [pem_pos=0, final_scores=0.5610314,
final_local_scores=0.4958518, pem=-0.32007957,
similarity=20.627563]
14: 18520327 Little_Rock,_Arkansas
    [pem_pos=0, final_scores=0.48119205,
final_local_scores=0.4377308, pem=-0.1562997,
similarity=18.002455]
```

Sample: b'216_RESEARCH'

RESEARCH ALERT - Royal Oak initiated . -- EVEREN Securities Inc said Friday it initiated coverage of Royal Oak Mines Inc with an outperform rating . -- It set earnings estimates of \$ 0.08 a share for fiscal 1996 , \$ 0.13 for 1997 , \$ 0.40 for 1998 and \$ 0.43 for 1999 . -- " Based on our simulated production , income and cash flow models , common shares of

Royal Oak are at a significant discount to the industry averages , " EVEREN said . -- The short-term price objective is \$ 5 a share and the long-term objective is \$ 9 . -- Royal Oak shares were down 1/16 at 3-11/16 . Reuters Chicago Newsdesk - 312-408-8787

Sample: b'217_SunGard'

[1 1 SunGard 1] to buy CheckFree unit . [2 2 SAN MATEO 2] , [3 3 Calif. 3] 1996-08-23 SunGard Shareholder Systems Inc , a subsidiary of SunGard Data Systems Inc , said it had entered into a definitive agreement to buy the Securities Products Business unit of CheckFree Corp . The company said in a statement on Friday the deal was expected to be finalized in September . Terms were not disclosed . The purchase is not expected to have a material effect o [4 4 SunGard 4] 's financial condition or results of operations . -- [5 5 New York 5] newsroom , (212) 859-1610
1: 1665223 SunGard
2: 108124 San_Mateo,_California
3: 5407 California
4: 1665223 SunGard
5: 645042 New_York_City
1: 1665223 SunGard
[pem_pos=0, final_scores=0.5671206,
final_local_scores=0.625111, pem=0.0, similarity=22.744272]
2: 108124 San_Mateo,_California
[pem_pos=0, final_scores=0.35635787,
final_local_scores=0.37394923, pem=-0.6619857,
similarity=18.849085]
3: 5407 California
[pem_pos=0, final_scores=0.5336891,
final_local_scores=0.57599354, pem=0.0, similarity=21.27851]
4: 1665223 SunGard
[pem_pos=0, final_scores=0.39521474,
final_local_scores=0.43732995, pem=0.0, similarity=17.14051]
5: 645042 New_York_City
[pem_pos=1, final_scores=0.38492006,
final_local_scores=0.40624448, pem=-1.7906687,
similarity=25.950804]

Sample: b'218_Alpha'

Alpha Techs closes Lockhart purchase . [1 1 NEW YORK 1] 1996-08-23 Alpha Technologies Group Inc said it had closed on its agreement to acquire Lockhart Industries Inc . The company said in a statement late on Thursday that it issued 280,556 shares of common stock for the stock of Lockhart . These shares are subject to post-closing adjustments . Lockhart , based in [2 2 Paramount 2] , [3 3 Calif. 3] , is a designer and manufacturer of sophisticated thermal management products . -- [4 4 New York 4] newsroom , (212) 859-1610
1: 645042 New_York_City

```
2: 107667 Paramount,_California
3: 5407 California
4: 645042 New_York_City
1: 645042 New_York_City
    [pem_pos=1, final_scores=0.46899885,
final_local_scores=0.42732626, pem=-1.7906687,
similarity=26.579916]
2: 107667 Paramount,_California
    [pem_pos=1, final_scores=0.48367882,
final_local_scores=0.54840416, pem=-1.7740386,
similarity=30.102684]
3: 5407 California
    [pem_pos=0, final_scores=0.6530399,
final_local_scores=0.6503163, pem=0.0, similarity=23.496449]
4: 645042 New_York_City
    [pem_pos=1, final_scores=0.4621808,
final_local_scores=0.4198786, pem=-1.7906687,
similarity=26.357662]
```

Sample: b'219_Analysts'

Analysts hold Dutch PTT estimates . [1 1 AMSTERDAM 1]
1996-08-23 [2 2 Dutch 2] post and telecoms group Koninklijke
PTT Nederland NV 's (Dutch PTT) first half results came
marginally below most analysts ' forecasts , giving scant
grounds to adjust full year estimates , analysts said on
Friday . PTT earlier announced an 8.5 percent rise in net
profit for the first six months of 1996 to 1.209 billion
guilders , a hair 's breadth below the 1.210-1.236 billion
forecast range . " They were pretty much in line with our
estimate of 1.210 billion , " said Peter Roe , at [3 3 Paribas
3] in [4 4 London 4] . " As we expected volume growth was
quite good ... there 's really no reason to change our
forecast and our largely positive view on the stock . " "
Coming just a million guilders under the forecast range is
n't overwhelmingly surprising , " said [5 5 ING 5] analyst
Steven Vrolijk , who is continuing to look for a nine to 10
percent rise in 1996 earnings . " They 've got a very sound
domestic business which is doing very well , we remain
positive on the stock , " said Roe . " We think it 's a solid
performer in the [6 6 Dutch 6] market . " Roe said he was
sticking by his forecast of a 2.45 billion guilder net for
1996 . Dutch PTT earlier on Friday repeated a forecast it
would improve on 1995 's 2.26 billion guilder net profit for
the whole year of 1996 . By 1403 [7 7 GMT 7] the shares were
down 1.70 guilders to 61.00 guilders , falling in a weaker [8
8 Amsterdam 8] bourse . -- Keiron Henderson , Amsterdam

Newsroom +31 20 504 5000

```
1: 844 Amsterdam
2: 21148 Netherlands
3: 564293 BNP_Paribas
4: 17867 London
5: 230965 ING_Group
```

```
6: 21148 Netherlands
7: 12701 Greenwich_Mean_Time
8: 844 Amsterdam
1: 844 Amsterdam
    [pem_pos=0, final_scores=0.5992191,
final_local_scores=0.5670014, pem=-0.0642803,
similarity=21.359732]
2: 21148 Netherlands
    [pem_pos=0, final_scores=0.61260337,
final_local_scores=0.58683604, pem=-0.7943452,
similarity=25.921844]
3: 564293 BNP_Paribas
    [pem_pos=0, final_scores=0.3980471,
final_local_scores=0.4614761, pem=0.0, similarity=17.86108]
4: 17867 London
    [pem_pos=0, final_scores=0.50630313,
final_local_scores=0.49540013, pem=-0.09086737,
similarity=19.367592]
5: 230965 ING_Group
    [pem_pos=0, final_scores=0.5401416,
final_local_scores=0.6016659, pem=0.0, similarity=22.044624]
6: 21148 Netherlands
    [pem_pos=0, final_scores=0.61900777,
final_local_scores=0.5938319, pem=-0.7943452,
similarity=26.130613]
7: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.6253446,
final_local_scores=0.7423052, pem=0.0, similarity=26.241581]
8: 844 Amsterdam
    [pem_pos=0, final_scores=0.5284486,
final_local_scores=0.48969537, pem=-0.0642803,
similarity=19.052767]
```

Sample: b'220_Boskalis'

```
[1 1 Boskalis 1] upgrades 1996 outlook . [ CORRECTED 13:12 [2
2 GMT 2] ] [3 3 PAPENDRECHT 3] , [4 4 Netherlands 4]
1996-08-23 Dredging group Koninklijke Boskalis Westminster NV
said on Friday it expected higher turnover and profits for
the second half of 1996 ( corrects from the full year
1996 ) . " An improvement is expected in capacity
utilisation , turnover and profits , " the company said in a
statement . The world 's largest dredger said in March that
it was uncertain whether it could hold 1996 full-year profit
steady at the previous year 's 70.9 million guilders , but
added long-term prospects were good . [5 5 Boskalis 5]
reported 1996 first-half net profit fell to 27.5 million
guilders from a year-earlier 41.4 million .
1: 15842094 Royal_Boskalis_Westminster
2: 12701 Greenwich_Mean_Time
3: 121173 Papendrecht
4: 21148 Netherlands
5: 15842094 Royal_Boskalis_Westminster
```

```
1: 15842094 Royal_Boskalis_Westminster
    [pem_pos=0, final_scores=0.49805713,
final_local_scores=0.5429163, pem=0.0, similarity=20.291418]
2: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.5051877,
final_local_scores=0.6527937, pem=0.0, similarity=23.57038]
3: 121173 Papendrecht
    [pem_pos=0, final_scores=0.4273513,
final_local_scores=0.60967994, pem=0.0, similarity=22.28378]
4: 21148 Netherlands
    [pem_pos=0, final_scores=0.4868542,
final_local_scores=0.57544905, pem=-0.24940795,
similarity=22.61858]
5: 15842094 Royal_Boskalis_Westminster
    [pem_pos=0, final_scores=0.39893675,
final_local_scores=0.43464231, pem=0.0, similarity=17.060307]
```

Sample: b'221_Greek'

```
[1 1 Greek 1] president dissolves parliament for snap vote .
[2 2 ATHENS 2] 1996-08-23 [3 3 Greek 3] President Costis
Stephnopoulos signed a decree on Thursday ordering the
dissolution of the 300-seat parliament ahead of snap
elections on September 22 . " The President of the Republic [4
4 Costis Stephanopoulos 4] signed the decree for the
dissolution of parliament , " a presidency statement said .
Socialist Prime Minister [5 5 Costas Simitis 5] announced the
snap poll on Thursday citing problems with the economy , the
country 's convergence with the [6 6 European Union 6] and
tense relations with neighbouring Turkey . Elections were
originally scheduled to be held in October next year .
```

```
1: 12108 Greece
2: 1216 Athens
3: 12108 Greece
4: 349084 Konstantinos_Stephanopoulos
5: 393824 Costas_Simitis
6: 9317 European_Union
1: 12108 Greece
    [pem_pos=1, final_scores=0.79144794,
final_local_scores=0.7825725, pem=-2.3247876,
similarity=40.085793]
2: 1216 Athens
    [pem_pos=0, final_scores=0.4511698,
final_local_scores=0.4472583, pem=-0.28831622,
similarity=19.004702]
3: 12108 Greece
    [pem_pos=1, final_scores=0.7573491,
final_local_scores=0.74532473, pem=-2.3247876,
similarity=38.974247]
4: 349084 Konstantinos_Stephanopoulos
    [pem_pos=0, final_scores=0.5815808,
final_local_scores=0.64596003, pem=0.0, similarity=23.36645]
5: 393824 Costas_Simitis
```

```
[pem_pos=0, final_scores=0.6147202,
final_local_scores=0.68566245, pem=0.0, similarity=24.551249]
6: 9317 European_Union
[pem_pos=0, final_scores=0.4670897,
final_local_scores=0.5107618, pem=-0.005814796,
similarity=19.363485]
```

Sample: b'222_Libyan'

```
[1 1 Libyan 1] man murdered in [2 2 Malta 2] . [3 3 VALLETTA
3] 1996-08-23 A [4 4 Libyan 4] man has been found stabbed to
death in [5 5 Malta 5] , police said on Friday . They said
the body of Amer Hishem Ali Mohammed , 23 , was found in a
pool of blood in [6 6 Sliema 6] , seven km ( four miles )
from [7 7 Valletta 7] , on Wednesday morning . He appeared to
have been killed on Tuesday night , suffering at least eight
stab wounds . Police Commissioner George Grech said police
were investigating the possibility that Mohammed had links to
an [8 8 Islamic 8] militant group as reported by the local
press . " What we know is that he was a fervant religious
man , we cannot exclude anything in investigations " he said .
1: 17633 Libya
2: 19137 Malta
3: 57664 Valletta
4: 17633 Libya
5: 19137 Malta
6: 1595031 Sliema
7: 57664 Valletta
8: 6037917 Islam
1: 17633 Libya
[pem_pos=0, final_scores=0.56610924,
final_local_scores=0.5661352, pem=-0.27370736,
similarity=22.47278]
2: 19137 Malta
[pem_pos=0, final_scores=0.5221821,
final_local_scores=0.47207326, pem=-0.39222723,
similarity=20.310314]
3: 57664 Valletta
[pem_pos=0, final_scores=0.46186846,
final_local_scores=0.438731, pem=-0.035661608,
similarity=17.376255]
4: 17633 Libya
[pem_pos=0, final_scores=0.5786763,
final_local_scores=0.57986283, pem=-0.27370736,
similarity=22.882439]
5: 19137 Malta
[pem_pos=0, final_scores=0.4846894,
final_local_scores=0.43111813, pem=-0.39222723,
similarity=19.08813]
6: 1595031 Sliema
[pem_pos=0, final_scores=0.24452119,
final_local_scores=0.25994354, pem=0.0, similarity=11.846947]
7: 57664 Valletta
```

```
[pem_pos=0, final_scores=0.3822282,
final_local_scores=0.35173607, pem=-0.035661608,
similarity=14.780152]
8: 6037917 Islam
[pem_pos=0, final_scores=0.27871865,
final_local_scores=0.33771205, pem=-0.21141872,
similarity=15.317442]
```

Sample: b'223_Deutsche'

```
[1 1 Deutsche Bahn 1] H1 pre-tax profit up 17.5 pct . [2 2 FR
ANKFURT 2] 1996-08-23 Six months to June 30 ( in millions of
marks unless stated ) Group pre-tax profit 188 vs 160 Group
sales 14,600 up 3.3 pct NOTE - Full name of the state-owned [3
3 German 3] railway company is [4 4 Deutsche Bahn AG 4] . The
company is earmarked for eventual privatisatio also covers
snap FAT8222 Revenue from long-distance passenger traffic
2,500 up 6.4 pct Revenue from commuter traffic 5,400 up 4.6
pct Revenue from freight traffic 3,200 down 5.1 pct Group
workforce on June 30 300,962 down 3.7 oct NOTE - Sales ,
profit compare with first half of 1995 , workforce compares
with Dec 31 . -- Frankfurt Newsroom , +49 69 756525
```

1: 172065 Deutsche_Bahn

2: 10992 Frankfurt

3: 11867 Germany

4: 172065 Deutsche_Bahn

1: 172065 Deutsche_Bahn

```
[pem_pos=0, final_scores=0.4331712,
final_local_scores=0.4119568, pem=0.0, similarity=16.383326]
```

2: 10992 Frankfurt

```
[pem_pos=0, final_scores=0.49881667,
final_local_scores=0.51115847, pem=-0.07062824,
similarity=19.727789]
```

3: 11867 Germany

```
[pem_pos=1, final_scores=0.6403703,
final_local_scores=0.687368, pem=-1.0465823,
similarity=30.293623]
```

4: 172065 Deutsche_Bahn

```
[pem_pos=0, final_scores=0.46508074,
final_local_scores=0.43890297, pem=0.0, similarity=17.187452]
```

Sample: b'224_Sea'

```
Sea lion paparazzi to keep tabs on whales . [1 1 LONDON 1]
1996-08-22 [2 2 U.S. 2] marine biologists have trained a pair
of sea lions to tag and photograph elusive whales as they
cruise through the Pacific depths , [3 3 New Scientist 3]
magazine reported on Thursday . James Harvey and Jennifer
Hurley of the [4 4 Moss Landing Marine Laboratories 4] in [5 5
California 5] say their sea lions , natural companions of
many species of whale , can go where no man or woman has ever
gone before . " Any diver knows that when a whale gets going
you ca n't keep up , " Harvey told the magazine . " That is
why we know only about five percent of what whales do . " The
```

sea lions -- 17-year-old Beaver and nine-year-old Sake -- have undergone six years of training for their mission . Beaver once worked for the [6 6 U.S. Navy 6] and Sake is an amusement park veteran . Harvey said they could accurately tag whales with a radio transmitter , and could also swim all the way around one of the giant mammals , filming it with a video camera . Their first assignment , later this year , will be documenting humpback whale migration off [7 7 Monterey 7] , [8 8 California 8] . The article did not spell out exactly how the sea lions manage to tag the whales but said in training they were taught to stick a radio transmitter on to a plastic model of a whale using suction cups .

```
1: 17867 London
2: 3434750 United_States
3: 39431 New_Scientist
4: 847620 Moss_Landing_Marine_Laboratories
5: 5407 California
6: 20518076 United_States_Navy
7: 40444 Monterey,_California
8: 5407 California
1: 17867 London
    [pem_pos=0, final_scores=0.63710636,
final_local_scores=0.66773134, pem=-0.09086737,
similarity=24.510298]
2: 3434750 United_States
    [pem_pos=0, final_scores=0.51947755,
final_local_scores=0.50636107, pem=-0.14075686,
similarity=19.965996]
3: 39431 New_Scientist
    [pem_pos=0, final_scores=0.31283534,
final_local_scores=0.40980786, pem=0.0, similarity=16.319197]
4: 847620 Moss_Landing_Marine_Laboratories
    [pem_pos=0, final_scores=0.3863197,
final_local_scores=0.50349706, pem=0.0, similarity=19.11507]
5: 5407 California
    [pem_pos=0, final_scores=0.6535885,
final_local_scores=0.6419553, pem=-0.08662026,
similarity=23.717995]
6: 20518076 United_States_Navy
    [pem_pos=0, final_scores=0.31067604,
final_local_scores=0.33418134, pem=0.0, similarity=14.06235]
7: 40444 Monterey,_California
    [pem_pos=0, final_scores=0.36036617,
final_local_scores=0.3446661, pem=-0.57613015,
similarity=17.508322]
8: 5407 California
    [pem_pos=0, final_scores=0.6593898,
final_local_scores=0.6482925, pem=-0.08662026,
similarity=23.907108]
```

Sample: b'225_RTRS'

RTRS - Australia Senate jeopardises rate cut - [1 1 Howard 1]

. [2 2 CANBERRA 2] 1996-08-23 [3 3 Australian 3] Prime Minister [4 4 John Howard 4] said the possibility of lower interest rates was being jeopardised by Parliament 's upper house where opposition parties planned to scrutinise the 1996/97 budget . " Every time the Senate hacks away at the budget , they hack away at the lower interest rate environment , " [5 5 Howard 5] told reporters on Thursday night after attending a [6 6 Liberal Party 6] function . Senior ministers have repeatedly warned since the fiscally-tight budget was handed down on Tuesday night that the chance of lower official rates could be hampered in the Senate . The budget contained sharp spending cuts in areas such as the labour market to reduce the deficit to about A\$ 5.6 billion . The conservative government 's plan for reform of the industrial relations environment and to partially sell [7 7 Telstra 7] has also been opposed by parties in the Senate such as the [8 8 Greens 8] and [9 9 Australian Democrats 9] as well as the official opposition , the [10 10 Labor Party 10] . Official cash rates were last cut on July 31 to 7.0 percent . -- Sydney Newsroom 61-2 373-1800

1: 4689709 John_Howard
2: 51983 Canberra
3: 4689264 Australia
4: 4689709 John_Howard
5: 4689709 John_Howard
6: 18453 Liberal_Party_of_Australia
7: 31532 Telstra
8: 171629 Australian_Greens
9: 1943 Australian_Democrats
10: 1495 Australian_Labor_Party
1: 4689709 John_Howard
[pem_pos=0, final_scores=1.057447,
final_local_scores=1.0568004, pem=0.0, similarity=35.626743]
2: 51983 Canberra
[pem_pos=0, final_scores=0.4656071,
final_local_scores=0.49222618, pem=-0.091910355,
similarity=19.278547]
3: 4689264 Australia
[pem_pos=0, final_scores=0.6537188,
final_local_scores=0.7229994, pem=-0.491009,
similarity=28.33564]
4: 4689709 John_Howard
[pem_pos=0, final_scores=0.627837,
final_local_scores=0.5787433, pem=-0.1997864,
similarity=22.447037]
5: 4689709 John_Howard
[pem_pos=0, final_scores=1.0416741,
final_local_scores=1.039571, pem=0.0, similarity=35.112587]
6: 18453 Liberal_Party_of_Australia
[pem_pos=1, final_scores=0.34224635,
final_local_scores=0.3117641, pem=-1.6500671,
similarity=22.366697]

```
7: 31532 Telstra
    [pem_pos=0, final_scores=0.36481112,
final_local_scores=0.42224383, pem=0.0, similarity=16.690311]
8: 171629 Australian_Greens
    [pem_pos=0, final_scores=0.39253396,
final_local_scores=0.3640296, pem=-1.0855342,
similarity=20.856386]
9: 1943 Australian_Democrats
    [pem_pos=0, final_scores=0.3796134,
final_local_scores=0.357446, pem=0.0, similarity=14.756615]
10: 1495 Australian_Labor_Party
    [pem_pos=0, final_scores=0.38153905,
final_local_scores=0.34165502, pem=-0.461849,
similarity=16.796986]
```

Sample: b'226_RTRS'

RTRS - [1 1 Toyota Australia 1] workers to return to work . [2 2 MELBOURNE 2] 1996-08-23 Over 2,000 striking workers voted on Friday to return to work on Monday at [3 3 Toyota Australia 3] 's [4 4 Melbourne 4] assembly line , ending a two-week stoppage .

```
1: 6764585 Toyota_Australia
2: 17306237 Melbourne
3: 6764585 Toyota_Australia
4: 17306237 Melbourne
1: 6764585 Toyota_Australia
    [pem_pos=0, final_scores=0.37557858,
final_local_scores=0.38311905, pem=0.0, similarity=15.522749]
2: 17306237 Melbourne
    [pem_pos=0, final_scores=0.48244441,
final_local_scores=0.49985343, pem=-0.16841523,
similarity=19.922205]
3: 6764585 Toyota_Australia
    [pem_pos=0, final_scores=0.44376916,
final_local_scores=0.45760685, pem=0.0, similarity=17.745615]
4: 17306237 Melbourne
    [pem_pos=0, final_scores=0.49048746,
final_local_scores=0.5086396, pem=-0.16841523,
similarity=20.1844]
```

Sample: b'227_RTRS'

RTRS - Niugini shares surge on bid talk . [1 1 SYDNEY 1] 1996-08-23 Shares in gold miner Niugini Mining Ltd surged 38 cents to A\$ 3.75 early on Friday following confirmation on Thursday from Battle Mountain Gold that it was considering acquiring the 49.6 percent of Niugini it did not already own . Niugini Mining Ltd said on Thursday that Battle Mountain had initiated talks about acquiring the shares in Niugini it does not already own . " Battle Mountain are set to take out the minorities there soon , " said a [2 2 Sydney 2] broker . Niugini holds copper and gold mining interests in [3 3 Australia 3] , [4 4 Chile 4] and [5 5 Papua New Guinea 5]

, where it has a 17.2 percent stake in the [6 6 Lihir 6] gold project . By 11.25 a.m. (0025 [7 7 GMT 7]) , Niugini Mining shares were at A\$ 3.65 , up 28 cents on turnover of 108,288 shares . -- Sydney Newsroom 61-2 9373 1800

1: 27862 Sydney
2: 27862 Sydney
3: 4689264 Australia
4: 5489 Chile
5: 22943 Papua_New_Guinea
6: 6368544 Lihir_Island
7: 12701 Greenwich_Mean_Time
1: 27862 Sydney
[pem_pos=0, final_scores=0.59184235,
final_local_scores=0.5922911, pem=-0.1443489,
similarity=22.549854]
2: 27862 Sydney
[pem_pos=0, final_scores=0.45614904,
final_local_scores=0.44406676, pem=-0.1443489,
similarity=18.126543]
3: 4689264 Australia
[pem_pos=0, final_scores=0.6678295,
final_local_scores=0.6472877, pem=-0.17168477,
similarity=24.33972]
4: 5489 Chile
[pem_pos=0, final_scores=0.69644004,
final_local_scores=0.7338092, pem=-0.27404457,
similarity=27.478338]
5: 22943 Papua_New_Guinea
[pem_pos=0, final_scores=0.38475853,
final_local_scores=0.39152855, pem=-0.26497513,
similarity=17.214682]
6: 6368544 Lihir_Island
[pem_pos=0, final_scores=0.46144974,
final_local_scores=0.614087, pem=-0.19961303,
similarity=23.50082]
7: 12701 Greenwich_Mean_Time
[pem_pos=0, final_scores=0.6505476,
final_local_scores=0.7693703, pem=0.0, similarity=27.049257]

Sample: b'228_South'
[1 1 South Korean 1] students throw " irreplaceable "
rocks . [2 2 SEOUL 2] 1996-08-23 Students at [3 3 South Korea
3] 's [4 4 Yonsei University 4] threw more than just ordinary
rocks at riot police -- some were samples that the geology
department had taken 30 years to collect , newspapers
reported on Friday . Geology professors were quoted as saying
that their collection of 10,000 rocks , gathered from across
the nation and abroad , were irreplaceable . " These are not
like missing window panes or broken desks . They are lost
forever , " said one professor . The students staged a
violent nine-day demonstration at the university to demand
unification with [5 5 North Korea 5] . Police ended the

```
protest on Tuesday after storming the campus .
1: 27019 South_Korea
2: 19159283 Seoul
3: 27019 South_Korea
4: 876553 Yonsei_University
5: 21255 North_Korea
1: 27019 South_Korea
    [pem_pos=0, final_scores=0.5703688,
final_local_scores=0.5142817, pem=0.0, similarity=19.436905]
2: 19159283 Seoul
    [pem_pos=0, final_scores=0.49427795,
final_local_scores=0.45634806, pem=-0.0378014,
similarity=17.91362]
3: 27019 South_Korea
    [pem_pos=0, final_scores=0.6069892,
final_local_scores=0.5542839, pem=-0.22449642,
similarity=21.851498]
4: 876553 Yonsei_University
    [pem_pos=0, final_scores=0.36378717,
final_local_scores=0.39071232, pem=0.0, similarity=15.749348]
5: 21255 North_Korea
    [pem_pos=0, final_scores=0.4059536,
final_local_scores=0.3896767, pem=-0.3090239,
similarity=17.398962]
```

Sample: b'229_S.'

```
[1 1 S. Korean 1] won ends up on dollar position unwinding .
[2 2 SEOUL 2] 1996-08-23 The won rose against the dollar on
Friday as banks unwound dollar positions on the belief that
the won would continue to strengthen , dealers said . The won
closed at 818.10 , after opening at 819.10 . It ranged from
817.60 to 819.30 . " The dollar is overbought at the moment ,
" said a western bank dealer . " The central bank 's
insistent intervention just above the 820 level convinced
players that it is serious about supporting the won . So some
foreign banks began unwinding their dollar positions . "
1: 27019 South_Korea
2: 19159283 Seoul
1: 27019 South_Korea
    [pem_pos=0, final_scores=0.504052,
final_local_scores=0.46929306, pem=0.0, similarity=18.094355]
2: 19159283 Seoul
    [pem_pos=0, final_scores=0.49794334,
final_local_scores=0.46262026, pem=-0.0378014,
similarity=18.100794]
```

Sample: b'230_Orii'

```
Orii - 96/97 group forecast . [1 1 TOKYO 1] 1996-08-23 Year
to May 31 , 1997 ( in billions of yen unless specified )
LATEST ACTUAL ( Group ) FORECAST YEAR-AGO Sales 8.70 8.67
Current prft 7 mln loss 371 mln Net nil loss 447 mln EPS nil
yen loss 48.61 yen NOTE - Orii Corp makes automation
```

equipment .

1: 30057 Tokyo

1: 30057 Tokyo

[pem_pos=0, final_scores=0.42743784,
final_local_scores=0.7394981, pem=-0.09283106,
similarity=26.66264]

Sample: b'231_Orii'

Orii - 95/96 group results . [1 1 TOKYO 1] 1996-08-23 Year
to May 31 , 1996 (Group) (in billions of yen unless
specified) Sales 8.67 vs 9.33 Operating loss 286 million vs
loss 48 million Current loss 371 million vs loss 278 million
Net loss 447 million vs loss 350 million EPS loss 48.61 yen
vs loss 38.11 yen Diluted EPS - vs - NOTE - Orii Corp makes
automation equipment .

1: 30057 Tokyo

1: 30057 Tokyo

[pem_pos=0, final_scores=0.39965433,
final_local_scores=0.7091488, pem=-0.09283106,
similarity=25.756958]

Sample: b'232_China'

[1 1 China 1] copper stocks owned by state reserve - trade .
Lynne O'Donnell [2 2 HONG KONG 2] 1996-08-23 Up to 100,000
tonnes of copper held in [3 3 Shanghai 3] bonded warehouses ,
confounding the world market as to its source and ultimate
fate , probably belongs to [4 4 China 4] 's strategic state
reserve , industry sources said on Friday . Around 40,000
tonnes of the copper have already been moved to warehouses
near the northern port of [5 5 Yingkou 5] , where some of the
strategic stockpile was stored , they said . Just who owns
the copper is a question that has kept traders and industry
analysts guessing since the metal was channelled into [6 6 Sh
anghai 6] by the China National Nonferrous Metals Import and
Export Corp (CNIEC) in June and July . It was unclear
whether or not the 40,000 tonnes had cleared customs -- which
would provide some concrete indication that the strategic
reserve , administered directly by the central government 's
State Planning Commission , owned the copper . Traders have
said the reserve could negotiate concessions on duties --
three percent import tax and 17 percent value-added tax --
that made the copper prohibitively expensive otherwise . But
one source , the head of a [7 7 Hong Kong 7] trading house ,
said it made no difference if the copper was customs cleared
or not . " If they spend all this money moving the copper to [
8 8 Yingkou 8] , it will be sitting there for years , " he
said . " Once it arrives in [9 9 Yingkou 9] , it is subject
to monitoring by the State Planning Commission , which has to
give permission for any more movement ; it is out of the
hands of traders , " he said . Mystery has surrounded the [10
10 Shanghai 10] stockpile in recent months , with traders
unsure not only of who owns it , but of its exact size and

what its owner planned to do with it . Trading sources generally agreed it would be cost-effective to take the copper back into a depleted central reserve as it had already served its purpose in taking advantage of long-term backwardation on the [11 11 London Metal Exchange 11] (LME) . A backwardation occurs when the spot price of a metal is higher than the forward price . CNIEC lent around 85,000 tonnes of copper onto the LME between April and June 1995 on behalf of the state reserve , running the state stockpile down to 115,000 tonnes from 200,000 tonnes previously . Traders in [12 12 Asia 12] said CNIEC could well have lent it to the market at around [13 13 US\$ 13] 2,700 a tonne , and then paid somewhere between \$ 2,200 and \$ 2,400 a tonne when it started taking the metal back earlier this year . This would have cleared CNIEC a healthy profit , which could then have been used to finance storage and other costs . Word that CNIEC had offered the copper to [14 14 European 14] trading houses in a series of secret meetings unnerved an already jittery market . Industry analysts Bloomsbury Minerals Economics (BME) said on Wednesday the motivation of the owners of the 85,000 tonnes , " whoever they are , is the most important short-term fundamental " in an already tight world market . BME repeated in its latest review rumours of involvement by Sumitomo Corp , with CNIEC said to be helping the [15 15 Japanese 15] trader unload its copper positions after it revealed in June losses of \$ 1.8 billion in a decade of unauthorised deals . [16 16 Sumitomo 16] and CNIEC have made no comments on the talk and [17 17 Chinese 17] traders said they know nothing of such an arrangement . Traders in [18 18 Shanghai 18] said on Thursday they were unaware of movements out of the [19 19 Shanghai 19] bonded warehouses . They reported more arrivals that were probably spot purchases . They also expressed concern that the tonnage in bonded warehouses would move onto the domestic market . But these concerns were irrelevant , a [20 20 Singapore 20] trader said , despite a forecast that domestic [21 21 Chinese 21] copper demand could hit one million tonnes this year . As with many commodities , " there is a desire (by the [22 22 Chinese 22] government) to keep a stockpile of the metal , " he said . " You do n't keep it to help industry , you keep it in case of emergency . " -- [23 23 Hong Kong 23] newsroom (852) 2843-6470

1: 5405 China

2: 13404 Hong_Kong

3: 27643 Shanghai

4: 5405 China

5: 1331885 Yingkou

6: 27643 Shanghai

7: 13404 Hong_Kong

8: 1331885 Yingkou

9: 1331885 Yingkou

10: 27643 Shanghai

```
11: 223968 London_Metal_Exchange
12: 689 Asia
13: 18717338 United_States_dollar
14: 9239 Europe
15: 15573 Japan
16: 1831418 The_Sumitomo_Bank
17: 5405 China
18: 27643 Shanghai
19: 27643 Shanghai
20: 27318 Singapore
21: 5405 China
22: 5405 China
23: 13404 Hong_Kong
1: 5405 China
    [pem_pos=0, final_scores=0.48697168,
final_local_scores=0.43733954, pem=-0.15669827,
similarity=17.992947]
2: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.6026353,
final_local_scores=0.5620947, pem=-0.22402541,
similarity=22.082026]
3: 27643 Shanghai
    [pem_pos=0, final_scores=0.5620036,
final_local_scores=0.5186347, pem=-0.10072064,
similarity=20.11454]
4: 5405 China
    [pem_pos=0, final_scores=0.49663073,
final_local_scores=0.44789058, pem=-0.15669827,
similarity=18.307812]
5: 1331885 Yingkou
    [pem_pos=0, final_scores=0.41779166,
final_local_scores=0.44609028, pem=0.0, similarity=17.401937]
6: 27643 Shanghai
    [pem_pos=0, final_scores=0.53119045,
final_local_scores=0.484976, pem=-0.10072064,
similarity=19.110098]
7: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.51112455,
final_local_scores=0.46213305, pem=-0.22402541,
similarity=19.09897]
8: 1331885 Yingkou
    [pem_pos=0, final_scores=0.53295773,
final_local_scores=0.5718918, pem=0.0, similarity=21.156105]
9: 1331885 Yingkou
    [pem_pos=0, final_scores=0.37951964,
final_local_scores=0.40428394, pem=0.0, similarity=16.154352]
10: 27643 Shanghai
    [pem_pos=0, final_scores=0.45816225,
final_local_scores=0.40520382, pem=-0.10072064,
similarity=16.729538]
11: 223968 London_Metal_Exchange
    [pem_pos=0, final_scores=0.4222322,
```

```
final_local_scores=0.5329656, pem=0.0, similarity=19.994469]
12: 689 Asia
    [pem_pos=0, final_scores=0.5204747,
final_local_scores=0.51861507, pem=-0.1494309,
similarity=20.378849]
13: 18717338 United_States_dollar
    [pem_pos=0, final_scores=0.4688331,
final_local_scores=0.5141364, pem=0.0, similarity=19.432568]
14: 9239 Europe
    [pem_pos=0, final_scores=0.36642712,
final_local_scores=0.3611967, pem=-0.25557294,
similarity=16.258389]
15: 15573 Japan
    [pem_pos=0, final_scores=0.8414125,
final_local_scores=0.8569687, pem=-0.99076414,
similarity=35.051296]
16: 1831418 The_Sumitomo_Bank
    [pem_pos=2, final_scores=0.25534993,
final_local_scores=0.39235777, pem=-1.6009561,
similarity=24.5047]
17: 5405 China
    [pem_pos=1, final_scores=0.6939338,
final_local_scores=0.6634142, pem=-1.4196125,
similarity=31.607391]
18: 27643 Shanghai
    [pem_pos=0, final_scores=0.45700026,
final_local_scores=0.40393448, pem=-0.10072064,
similarity=16.691658]
19: 27643 Shanghai
    [pem_pos=0, final_scores=0.63893783,
final_local_scores=0.60267365, pem=-0.10072064,
similarity=22.62243]
20: 27318 Singapore
    [pem_pos=0, final_scores=0.50013804,
final_local_scores=0.46841615, pem=-0.1922004,
similarity=19.113401]
21: 5405 China
    [pem_pos=1, final_scores=0.79151803,
final_local_scores=0.7700102, pem=-1.4196125,
similarity=34.788425]
22: 5405 China
    [pem_pos=1, final_scores=0.70839065,
final_local_scores=0.6792062, pem=-1.4196125,
similarity=32.078655]
23: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.48810118,
final_local_scores=0.43698353, pem=-0.22402541,
similarity=18.34846]
```

Sample: b'233_Companion'
Companion Marble posts 1st final result . [1 1 HONG KONG 1]
1996-08-23 Year ended March 31 (in million [2 2 HK\$ 2]

unless stated) Shr ([3 3 H.K. 3] cents) 14.0 Dividend ([4 4 H.K. 4] cents) nil Exceptional items nil Net 56.06 Turnover 531.52 Company name Companion Marble ([5 5 Holdings 5]) [6 6 Ltd 6] Books close N / A Dividend payable N / A NOTE - Marble and granite products distributor Companion Marble , a spinoff of construction materials concern Companion Building Material ([7 7 Holdings 7]) [8 8 Ltd 8] , was listed on the Stock Exchange on April 25 , 1996 . -- Hong Kong News Room (852) 2843 6368

1: 13404 Hong_Kong
2: 309734 Hong_Kong_dollar
3: 13404 Hong_Kong
4: 13404 Hong_Kong
5: 146080 Holding_company
6: 538325 Limited_company
7: 146080 Holding_company
8: 538325 Limited_company

1: 13404 Hong_Kong
[pem_pos=0, final_scores=0.53851706,
final_local_scores=0.5221407, pem=-0.22402541,
similarity=20.889717]
2: 309734 Hong_Kong_dollar
[pem_pos=0, final_scores=0.55944103,
final_local_scores=0.62256074, pem=0.0, similarity=22.668167]
3: 13404 Hong_Kong
[pem_pos=0, final_scores=0.6507351,
final_local_scores=0.6447218, pem=-0.6931472,
similarity=27.09894]
4: 13404 Hong_Kong
[pem_pos=0, final_scores=0.6820497,
final_local_scores=0.6789282, pem=-0.6931472,
similarity=28.119726]
5: 146080 Holding_company
[pem_pos=13, final_scores=0.49267882,
final_local_scores=0.50023156, pem=-2.4849067,
similarity=32.53093]
6: 538325 Limited_company
[pem_pos=0, final_scores=0.4376592,
final_local_scores=0.4453559, pem=-0.403316,
similarity=19.573317]
7: 146080 Holding_company
[pem_pos=13, final_scores=0.61605155,
final_local_scores=0.63499755, pem=-2.4849067,
similarity=36.552616]
8: 538325 Limited_company
[pem_pos=0, final_scores=0.4867578,
final_local_scores=0.49898863, pem=-0.403316,
similarity=21.173824]

Sample: b'234_Softbank'
[1 1 Softbank 1] to procure \$ 900 mln via forex by Sept 5 . [
2 2 TOKYO 2] 1996-08-23 [3 3 Softbank Corp 3] said on Friday

that it would procure \$ 900 million through the foreign exchange market by September 5 as part of its acquisition of [4 4 U.S. 4] firm , Kingston Technology Co . " It is in the contract that we pay ([5 5 Kingston 5]) \$ 900 million by September 5 , " he said , adding that [6 6 Softbank 6] had already started making forward transactions to buy dollars . On August 15 , computer software retailer [7 7 Softbank 7] said it would buy 80 percent of [8 8 Kingston 8] , the world 's largest maker of memory boards , for about \$ 1.5 billion in the latest in a series of high-profile acquisitions it has made in the [9 9 United States 9] .

1: 370229 SoftBank
2: 30057 Tokyo
3: 370229 SoftBank
4: 3434750 United_States
5: 4687132 Kingston_Technology
6: 370229 SoftBank
7: 370229 SoftBank
8: 4687132 Kingston_Technology
9: 3434750 United_States
1: 370229 SoftBank
[pem_pos=0, final_scores=0.46844023,
final_local_scores=0.43853623, pem=0.0, similarity=17.176508]
2: 30057 Tokyo
[pem_pos=0, final_scores=0.66486067,
final_local_scores=0.7437586, pem=-0.09283106,
similarity=26.789783]
3: 370229 SoftBank
[pem_pos=0, final_scores=0.5300725,
final_local_scores=0.50586015, pem=0.0, similarity=19.185589]
4: 3434750 United_States
[pem_pos=0, final_scores=0.4279496,
final_local_scores=0.47950965, pem=-0.14075686,
similarity=19.164696]
5: 4687132 Kingston_Technology
[pem_pos=15, final_scores=0.80304784,
final_local_scores=0.8841614, pem=-4.106293,
similarity=52.80551]
6: 370229 SoftBank
[pem_pos=0, final_scores=0.4288628,
final_local_scores=0.3953039, pem=0.0, similarity=15.886371]
7: 370229 SoftBank
[pem_pos=0, final_scores=0.4891199,
final_local_scores=0.4611256, pem=0.0, similarity=17.85062]
8: 4687132 Kingston_Technology
[pem_pos=15, final_scores=0.7924275,
final_local_scores=0.8725603, pem=-4.106293,
similarity=52.459312]
9: 3434750 United_States
[pem_pos=0, final_scores=0.4665566,
final_local_scores=0.52168185, pem=-0.073457584,
similarity=20.057215]

Sample: b'235_Shanghai'

Shanghai Post and Telecomm net down . [1 1 SHANGHAI 1] 1996-08-23 Half-year ended June 30 , 1996 (in millions of yuan unless stated) Turnover 115.259 vs 123.157 Net profit 20.318 vs 22.828 Net asset per share 3.02 yuan (no comparative figure) Earnings per share 0.14 yuan (no comparative figure) Company name : Shanghai Posts and Telecommunications Equipment Co Note : the figures were unaudited .

1: 27643 Shanghai

1: 27643 Shanghai

[pem_pos=0, final_scores=0.40881598, final_local_scores=0.71915656, pem=-0.10072064, similarity=26.098513]

Sample: b'236_Promodes'

Promodes set to sell [1 1 German 1] assets - paper . [2 2 PA RIS 2] 1996-08-23 [3 3 German 3] retailer Promodes is in advanced talks about selling its [4 4 German 4] assets and the retailer 's board might decide as soon as Tuesday to sell the assets to Spar AG , [5 5 Les Echos 5] newspaper said on Friday . It said investment bank Rothschild & Cie was an intermediary in the talks and added that unlisted [6 6 German 6] retailers [7 7 Metro 7] , Rewe and [8 8 Lidl 8] were also still in discussions . Promodes has in [9 9 Germany 9] its Promo hypermarket unit with 36 Continent superstores , which is 1995 generated 4.7 percent of total Promodes sales . The [1 0 10 French 10] group entered the [11 11 German 11] market in 1990 , buying In June , Promodes signed an outline agreement to sell its Specia unit -- which runs 100 [12 12 Dia 12] stores in [13 13 France 13] -- to [14 14 Germany 14] 's [15 15 Aldi 15] . Promodes was not immediately available for comment . -- [16 16 Paris 16] newsroom +33 1 4221 5452

1: 11867 Germany

2: 22989 Paris

3: 11867 Germany

4: 11867 Germany

5: 2688005 Les_Echos_(France)

6: 11867 Germany

7: 305874 Metro_AG

8: 21470795 Lidl

9: 11867 Germany

10: 5843419 France

11: 11867 Germany

12: 8240811 Dia_(supermarket_chain)

13: 5843419 France

14: 11867 Germany

15: 404177 Aldi

16: 22989 Paris

1: 11867 Germany

[pem_pos=1, final_scores=0.8738738,

```
final_local_scores=0.85264283, pem=-1.0465823,
similarity=35.225754]
2: 22989 Paris
    [pem_pos=0, final_scores=0.53443825,
final_local_scores=0.5296808, pem=-0.115301244,
similarity=20.52347]
3: 11867 Germany
    [pem_pos=1, final_scores=0.87631756,
final_local_scores=0.85531217, pem=-1.0465823,
similarity=35.305412]
4: 11867 Germany
    [pem_pos=1, final_scores=0.6667682,
final_local_scores=0.6264114, pem=-1.0465823,
similarity=28.474556]
5: 2688005 Les_Echos_(France)
    [pem_pos=0, final_scores=0.35770005,
final_local_scores=0.48347235, pem=0.0, similarity=18.517492]
6: 11867 Germany
    [pem_pos=1, final_scores=0.79655534,
final_local_scores=0.7681841, pem=-1.0465823,
similarity=32.705338]
7: 305874 Metro_AG
    [pem_pos=18, final_scores=0.5373858,
final_local_scores=0.6559021, pem=-4.16861,
similarity=46.332687]
8: 21470795 Lidl
    [pem_pos=0, final_scores=0.3710512,
final_local_scores=0.45510513, pem=0.0, similarity=17.670958]
9: 11867 Germany
    [pem_pos=0, final_scores=0.8015931,
final_local_scores=0.7736871, pem=-0.15529114,
similarity=28.022577]
10: 5843419 France
    [pem_pos=0, final_scores=0.63651115,
final_local_scores=0.61922586, pem=-0.80223274,
similarity=26.931314]
11: 11867 Germany
    [pem_pos=1, final_scores=0.8297685,
final_local_scores=0.8044644, pem=-1.0465823,
similarity=33.788013]
12: 8240811 Dia_(supermarket_chain)
    [pem_pos=4, final_scores=0.7629666,
final_local_scores=0.93193454, pem=-2.6866455,
similarity=46.510895]
13: 5843419 France
    [pem_pos=0, final_scores=0.6555843,
final_local_scores=0.6400603, pem=-0.16240834,
similarity=24.07359]
14: 11867 Germany
    [pem_pos=0, final_scores=0.96805173,
final_local_scores=0.9555178, pem=-0.15529114,
similarity=33.44877]
```

```
15: 404177 Aldi
    [pem_pos=0, final_scores=0.34097815,
final_local_scores=0.41474187, pem=0.0, similarity=16.466438]
16: 22989 Paris
    [pem_pos=0, final_scores=0.5900454,
final_local_scores=0.59042317, pem=-0.115301244,
similarity=22.336143]
```

Sample: b'237_PRESS'

PRESS DIGEST - [1 1 Ireland 1] - August 23 . [2 2 DUBLIN 2]
1996-08-23 Following are highlights of stories in the [3 3 Ir
ish 3] press on Friday . [4 4 Reuters 4] has not verified
these stories and does not vouch for their accuracy . [5 5 IR
ISH INDEPENDENT 5] - [6 6 Ireland 6] 's biggest mortgage
lender Irish Permanent ended weeks of stalemate when it
announced it was increasing its mortgage lending rate by a
quarter of a percentage point . - Two investors who claim to
be owed nearly one million [7 7 Irish 7] pounds by fund
manager Tony Taylor believe they may have lost their money .
- A second [8 8 Japanese 8] trawler was under arrest on
Thursday night as the Irish Navy and Air Corps continued a
cat and mouse game with up to 40 vessels off the [9 9 Irish 9]
coast . - The Irish Department of Enterprise and Employment
has widened its probe into Taylor Asset Managers to include
the investigation of investments of 10 more investors . - [10
10 Irish 10] exploration company Ivernia and its [11 11 South
African 11] partner Minorco have received planning permission
from the local county council for a major lead and zinc mine
at Lisheen , [12 12 County Tipperary 12] . - Building
materials firm [13 13 CRH 13] refused to comment on reports
that it is about to pay 180 million pounds sterling for [14 14
U.S. 14] stone and concrete business Tilcon Inc . [15 15 IRISH
TIMES 15] - Mortgage lending rates are on the way up with
banks and building societies poised to add around a quarter
of a percentage point to their main variable rates of
interest . - Members of a [16 16 County Antrim 16] Protestant
family who were driven into exile by loyalist
paramilitaries two years ago returned yesterday to live in [17
17 Northern Ireland 17] in defiance of the threat hanging
over them . - Talks will resume next Tuesday in an attempt to
avoid a major strike in [18 18 Irish 18] retail chain [19 19 D
unnes Stores 19] . - Top dealers in the [20 20 London 20]
office of [21 21 U.S. 21] bankers [22 22 Merrill Lynch 22]
are transferring to [23 23 Dublin 23] . - The [24 24 Irish 24]
plastics industry called on the government to support
incineration to deal with plastics not suitable for
recycling .

```
1: 14560 Republic_of_Ireland
2: 8504 Dublin
3: 14560 Republic_of_Ireland
4: 18998750 Reuters
5: 169544 Irish_Independent
```

```
6: 14560 Republic_of_Ireland
7: 14560 Republic_of_Ireland
8: 15573 Japan
9: 14560 Republic_of_Ireland
10: 14560 Republic_of_Ireland
11: 17416221 South_Africa
12: 144087 County_Tipperary
13: 1003019 CRH_plc
14: 3434750 United_States
15: 147378 The_Irish_Times
16: 46761 County_Antrim
17: 21265 Northern_Ireland
18: 14560 Republic_of_Ireland
19: 958473 Dunnes_Stores
20: 17867 London
21: 3434750 United_States
22: 33731299 Merrill_Lynch
23: 8504 Dublin
24: 14560 Republic_of_Ireland
1: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.569482,
final_local_scores=0.5167958, pem=-1.5040284,
similarity=27.691072]
2: 8504 Dublin
    [pem_pos=0, final_scores=0.4892456,
final_local_scores=0.44565028, pem=-0.23150389,
similarity=18.64776]
3: 14560 Republic_of_Ireland
    [pem_pos=3, final_scores=0.60279745,
final_local_scores=0.5531879, pem=-2.2576945,
similarity=32.875637]
4: 18998750 Reuters
    [pem_pos=0, final_scores=0.283701,
final_local_scores=0.3401566, pem=0.0, similarity=14.2406645]
5: 169544 Irish_Independent
    [pem_pos=0, final_scores=0.54920834,
final_local_scores=0.60995746, pem=0.0, similarity=22.29206]
6: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.5126135,
final_local_scores=0.4546756, pem=-1.5040284,
similarity=25.83728]
7: 14560 Republic_of_Ireland
    [pem_pos=3, final_scores=0.5975494,
final_local_scores=0.54745513, pem=-2.2576945,
similarity=32.70456]
8: 15573 Japan
    [pem_pos=0, final_scores=0.77026945,
final_local_scores=0.84795135, pem=-0.99076414,
similarity=34.7822]
9: 14560 Republic_of_Ireland
    [pem_pos=3, final_scores=0.5291441,
final_local_scores=0.47273284, pem=-2.2576945,
```

```
similarity=30.474695]
10: 14560 Republic_of_Ireland
    [pem_pos=3, final_scores=0.55757076,
final_local_scores=0.5037846, pem=-2.2576945,
similarity=31.401344]
11: 17416221 South_Africa
    [pem_pos=0, final_scores=0.3875975,
final_local_scores=0.42455167, pem=-0.1389451,
similarity=17.514786]
12: 144087 County_Tipperary
    [pem_pos=0, final_scores=0.3198654,
final_local_scores=0.29820177, pem=0.0, similarity=12.988648]
13: 1003019 CRH_plc
    [pem_pos=2, final_scores=0.54699975,
final_local_scores=0.6433357, pem=-1.4554272,
similarity=31.202972]
14: 3434750 United_States
    [pem_pos=0, final_scores=0.5005566,
final_local_scores=0.49497896, pem=-0.14075686,
similarity=19.62633]
15: 147378 The_Irish_Times
    [pem_pos=0, final_scores=0.45935112,
final_local_scores=0.4901262, pem=0.0, similarity=18.716057]
16: 46761 County_Antrim
    [pem_pos=0, final_scores=0.43616742,
final_local_scores=0.4230029, pem=0.0, similarity=16.712963]
17: 21265 Northern_Ireland
    [pem_pos=0, final_scores=0.41521138,
final_local_scores=0.3585832, pem=-0.25044483,
similarity=16.15251]
18: 14560 Republic_of_Ireland
    [pem_pos=3, final_scores=0.64030653,
final_local_scores=0.59416085, pem=-2.2576945,
similarity=34.09835]
19: 958473 Dunnes_Stores
    [pem_pos=0, final_scores=0.5304292,
final_local_scores=0.6107178, pem=0.0, similarity=22.314749]
20: 17867 London
    [pem_pos=0, final_scores=0.46739805,
final_local_scores=0.46724182, pem=-0.09086737,
similarity=18.52729]
21: 3434750 United_States
    [pem_pos=0, final_scores=0.5039049,
final_local_scores=0.49863642, pem=-0.14075686,
similarity=19.735477]
22: 33731299 Merrill_Lynch
    [pem_pos=0, final_scores=0.4718604,
final_local_scores=0.57726216, pem=0.0, similarity=21.316368]
23: 8504 Dublin
    [pem_pos=0, final_scores=0.40740138,
final_local_scores=0.3562479, pem=-0.23150389,
similarity=15.9798155]
```

24: 14560 Republic_of_Ireland
[pem_pos=3, final_scores=0.5108282,
final_local_scores=0.45272547, pem=-2.2576945,
similarity=29.877636]

Sample: b'238_Ford'
Ford China JV posts 77 percent net drop in H1 96 . [1 1 SHAN
GHAI 1] 1996-08-24 A [2 2 Chinese 2] truck maker in which [3 3
Ford Motor Co 3] has a 20 percent stake said it posted a 77
percent drop in post-tax profits in the first half of 1996 .
Jiangling Motors Corp , in a statement in Saturday 's edition
of the China Securities newspaper , said net profit in the
period was 3.385 million yuan , down from 14.956 million in
the same 1995 period . Turnover fell to 937.891 million yuan
from 1.215 billion , while net assets per share were 1.88
yuan , unchanged , and earnings per share fell to 0.005 yuan
from 0.02 yuan , the statement said . In the first half , the
company said it produced 8,333 vehicles and sold 9,018 , but
it did not explain the difference . It blamed the drop in
profits on a weak vehicle market and said as its engine plant
had only just started trial production , the company 's
results would not improve in the short-term . [4 4 Ford 4]
owns 138.643 million shares in the firm . The majority
shareholder , with 51 percent , or 353.24 million shares , is
Jiangling Motors Group . The company , in the southern
province of [5 5 Jiangxi 5] , had about eight percent of [6 6
China 6] 's light truck market in 1994 . (\$ 1 = 8.3 yuan)

1: 27643 Shanghai
2: 5405 China
3: 30433662 Ford_Motor_Company
4: 30433662 Ford_Motor_Company
5: 173817 Jiangxi
6: 5405 China
1: 27643 Shanghai
[pem_pos=0, final_scores=0.627446,
final_local_scores=0.6342318, pem=-0.10072064,
similarity=23.56419]
2: 5405 China
[pem_pos=1, final_scores=0.5671057,
final_local_scores=0.5499459, pem=-1.4196125,
similarity=28.221268]
3: 30433662 Ford_Motor_Company
[pem_pos=0, final_scores=0.27223188,
final_local_scores=0.2938911, pem=0.0, similarity=12.860009]
4: 30433662 Ford_Motor_Company
[pem_pos=0, final_scores=0.49552196,
final_local_scores=0.53780156, pem=-0.3393453,
similarity=21.984198]
5: 173817 Jiangxi
[pem_pos=0, final_scores=0.31296486,
final_local_scores=0.33853206, pem=0.0, similarity=14.192184]
6: 5405 China

```
[pem_pos=0, final_scores=0.50484365,
final_local_scores=0.48193407, pem=-0.15669827,
similarity=19.323736]

Sample: b'239_TENNIS'
TENNIS - RESULTS AT HAMLET CUP . [1 1 COMMACK 1] , [1 1 New York 1] 1996-08-24 Results from the Waldbaum Hamlet Cup tennis tournament on Saturday ( prefix number denotes seeding ) :
Semifinals : [2 2 Martin Damm 2] ( [3 3 Czech Republic 3] )
beat [4 4 Adrian Voinea 4] ( [5 5 Romania 5] ) 5-7 7-5 7-5 5
- [6 6 Andrey Medvedev 6] ( [7 7 Ukraine 7] ) beat [8 8 Karol Kucera 8] ( [9 9 Slovakia 9] ) 7-6 ( 7-0 ) 6-3
1: 127266 Commack,_New_York
2: 4026614 Martin_Damm
3: 5321 Czech_Republic
4: 4036370 Adrian_Voinea
5: 25445 Romania
6: 702849 Andriy_Medvedev
7: 31750 Ukraine
8: 3039956 Karol_Kučera
9: 26830 Slovakia
1: 645042 New_York_City
    [winner: 8210131 New_York, pem_pos=0,
final_scores=0.15654163, final_local_scores=0.18355078,
pem=-0.2889045, similarity=11.138342]
    gt: 645042 New_York_City, pem_pos=1,
final_scores=0.11133908, final_local_scores=0.14087015,
pem=-1.7906687, similarity=18.031494 ]
1: 127266 Commack,_New_York
    [pem_pos=0, final_scores=0.48698074,
final_local_scores=0.6218137, pem=0.0, similarity=22.645874]
2: 4026614 Martin_Damm
    [pem_pos=0, final_scores=0.8554134,
final_local_scores=1.0001144, pem=0.0, similarity=33.93512]
3: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.5540819,
final_local_scores=0.5522134, pem=-0.27794555,
similarity=22.080372]
4: 4036370 Adrian_Voinea
    [pem_pos=0, final_scores=0.60093176,
final_local_scores=0.74414945, pem=0.0, similarity=26.296618]
5: 25445 Romania
    [pem_pos=0, final_scores=0.7175024,
final_local_scores=0.74614537, pem=-0.21241479,
similarity=27.511324]
6: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=0.91259533,
final_local_scores=1.0405068, pem=0.0, similarity=35.140514]
7: 31750 Ukraine
    [pem_pos=0, final_scores=0.5907418,
final_local_scores=0.617002, pem=-0.35407752,
similarity=24.427813]
```

```
8: 3039956 Karol_Kučera
    [pem_pos=0, final_scores=0.75541025,
final_local_scores=0.877204, pem=0.0, similarity=30.26723]
9: 26830 Slovakia
    [pem_pos=0, final_scores=0.65025383,
final_local_scores=0.6559189, pem=-0.36602598,
similarity=25.654148]
1: 8210131 New_York
    [pem_pos=0, final_scores=0.15654163,
final_local_scores=0.18355078, pem=-0.2889045,
similarity=11.138342]

Sample: b'240_TENNIS'
TENNIS - RESULTS AT [1 1 TOSHIBA CLASSIC 1] . [2 2 CARLSBAD 2
] , [3 3 Calif. 3] 1996-08-24 Results from the $ 450,000 [4 4
Toshiba Classic 4] tennis tournament on Saturday ( prefix
number denotes seeding ) : Semifinals : 1 - [5 5 Arantxa Sanc
hez Vicario 5] ( [6 6 Spain 6] ) beat 3 - [7 7 Jana Novotna 7]
( [8 8 Czech Republic 8] ) 1-6 , 6-2 6-3 4 - [9 9 Kimiko Date
9] ( [10 10 Japan 10] ) beat 2 - [11 11 Conchita Martinez 11]
( [12 12 Spain 12] ) 6-2 7-5 .
1: 31726664 Southern_California_Open
2: 108030 Carlsbad,_California
3: 5407 California
4: 31726664 Southern_California_Open
5: 534624 Arantxa_Sánchez_Vicario
6: 26667 Spain
7: 307050 Jana_Novotná
8: 5321 Czech_Republic
9: 495296 Kimiko_Date-Krumm
10: 15573 Japan
11: 784152 Conchita_Martínez
12: 26667 Spain
1: 31726664 Southern_California_Open
    [pem_pos=0, final_scores=0.46346766,
final_local_scores=0.4913178, pem=0.0, similarity=18.751616]
2: 108030 Carlsbad,_California
    [pem_pos=0, final_scores=0.48723322,
final_local_scores=0.54156035, pem=-0.36353475,
similarity=22.227913]
3: 5407 California
    [pem_pos=0, final_scores=0.48923755,
final_local_scores=0.56321263, pem=0.0, similarity=20.8971]
4: 31726664 Southern_California_Open
    [pem_pos=0, final_scores=0.49974817,
final_local_scores=0.53094876, pem=0.0, similarity=19.934282]
5: 534624 Arantxa_Sánchez_Vicario
    [pem_pos=0, final_scores=0.6760785,
final_local_scores=0.7318282, pem=0.0, similarity=25.928926]
6: 26667 Spain
    [pem_pos=0, final_scores=0.6355676,
final_local_scores=0.6833548, pem=-0.18857051,
```

```
similarity=25.507858]
7: 307050 Jana_Novotná
    [pem_pos=0, final_scores=0.8800729,
final_local_scores=0.96876377, pem=0.0, similarity=32.999557]
8: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.5864671,
final_local_scores=0.64822197, pem=-0.27794555,
similarity=24.945461]
9: 495296 Kimiko_Date-Krumm
    [pem_pos=0, final_scores=0.7483358,
final_local_scores=0.81703275, pem=0.0, similarity=28.4716]
10: 15573 Japan
    [pem_pos=0, final_scores=0.659122,
final_local_scores=0.7364331, pem=-0.13549899,
similarity=26.80321]
11: 784152 Conchita_Martínez
    [pem_pos=0, final_scores=0.83356637,
final_local_scores=0.913781, pem=0.0, similarity=31.358759]
12: 26667 Spain
    [pem_pos=0, final_scores=0.6042389,
final_local_scores=0.64913297, pem=-0.18857051,
similarity=24.48661]
```

Sample: b'241_RALLYING'

RALLYING - [1 1 KANKKUNEN 1] IN COMMAND AS [2 2 MCRAE 2]
ROLLS OUT . [3 3 JYVASKYLA 3] , [4 4 Finland 4] 1996-08-24 [5
5 Finland 5] 's [6 6 Juha Kankkunen 6] produced an impressive
performance in his [7 7 Toyota 7] on Saturday to open up a 37
seconds lead after six stages of the 1,000 Lakes Rally ,
sixth round of the world championship . On a weekend
overshadowed by Friday 's fatal accident , four times world
champion Kankunnen emerged from the first five of Saturday 's
10 stages with a commanding advantage over his country 's
latest prospect , [8 8 Marcus Gronholm 8] , also in a [9 9 To
yota 9] . World championship leader Tommi Makinen in his [10 1
0 Mitsubishi 10] was third but current world champion [11 11 C
olin McRae 11] ended a bad week by crashing out . After being
fined \$ 250,000 by the sports governing body on Tuesday , the
[12 12 British 12] driver rolled his [13 13 Subaru 13] 6.5 km
into stage six . He and co-driver [14 14 Derek Ringer 14]
were unhurt but team boss [15 15 David Richards 15] was
furious with them . " It 's not unfortunate , it 's
incompetent , " he declared . [16 16 Kankkunen 16] has set an
astonishing pace for a driver who has not rallied for three
months . " If you do a lot of something , sometimes it 's
good to have a break . It 's not bad for an old man ! " said
the 37-year-old veteran . [17 17 Ford 17] had a poor morning
with [18 18 Spaniard 18] [19 19 Carlos Sainz 19] losing 90
seconds through turbo trouble while [20 20 Belgian 20] [21 21
Bruno Thiry 21] dropped four minutes when a transmission
shaft snapped .

1: 412742 Juha_Kankkunen

```
2: 1109927 Colin_McRae
3: 258746 Jyväskylä
4: 10577 Finland
5: 10577 Finland
6: 412742 Juha_Kankunen
7: 30984 Toyota
8: 944600 Marcus_Grönholm
9: 30984 Toyota
10: 203872 Mitsubishi_Motors
11: 1109927 Colin_McRae
12: 31717 United_Kingdom
13: 23901899 Subaru
14: 8020810 Derek_Ringer
15: 1085647 David_Richards_(racing)
16: 412742 Juha_Kankunen
17: 11175992 Ford_World_Rally_Team
18: 26667 Spain
19: 407139 Carlos_Sainz
20: 3343 Belgium
21: 8958042 Bruno_Thiry
1: 412742 Juha_Kankunen
    [pem_pos=0, final_scores=0.4189257,
final_local_scores=0.40929812, pem=0.0, similarity=16.303986]
2: 1109927 Colin_McRae
    [pem_pos=0, final_scores=0.93181795,
final_local_scores=0.9703441, pem=0.0, similarity=33.04672]
3: 258746 Jyväskylä
    [pem_pos=0, final_scores=0.47093314,
final_local_scores=0.56293875, pem=0.0, similarity=20.888927]
4: 10577 Finland
    [pem_pos=0, final_scores=0.7722139,
final_local_scores=0.84339523, pem=-0.2445894,
similarity=30.588425]
5: 10577 Finland
    [pem_pos=0, final_scores=0.6269787,
final_local_scores=0.68474793, pem=-0.2445894,
similarity=25.854073]
6: 412742 Juha_Kankunen
    [pem_pos=0, final_scores=0.75958437,
final_local_scores=0.781416, pem=0.0, similarity=27.408726]
7: 30984 Toyota
    [pem_pos=0, final_scores=0.4789399,
final_local_scores=0.4855715, pem=-0.17762902,
similarity=19.546108]
8: 944600 Marcus_Grönholm
    [pem_pos=0, final_scores=0.68315727,
final_local_scores=0.6930344, pem=0.0, similarity=24.771242]
9: 30984 Toyota
    [pem_pos=0, final_scores=0.48196805,
final_local_scores=0.48887932, pem=-0.17762902,
similarity=19.644821]
10: 203872 Mitsubishi_Motors
```

```
[pem_pos=1, final_scores=0.39211494,
final_local_scores=0.4161077, pem=-1.328055,
similarity=23.729366]
11: 1109927 Colin_McRae
    [pem_pos=0, final_scores=0.6382186,
final_local_scores=0.64062417, pem=0.0, similarity=23.207216]
12: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.4300276,
final_local_scores=0.49801022, pem=-0.48936236,
similarity=21.61256]
13: 23901899 Subaru
    [pem_pos=0, final_scores=0.35804933,
final_local_scores=0.36551982, pem=0.0, similarity=14.997553]
14: 8020810 Derek_Ringer
    [pem_pos=0, final_scores=0.40265614,
final_local_scores=0.48679966, pem=0.0, similarity=18.616785]
15: 1085647 David_Richards_(racing)
    [pem_pos=0, final_scores=0.30478054,
final_local_scores=0.32110325, pem=-0.3921967,
similarity=15.8049]
16: 412742 Juha_Kankkunen
    [pem_pos=0, final_scores=0.5572176,
final_local_scores=0.560361, pem=0.0, similarity=20.812004]
17: 11175992 Ford_World_Rally_Team
    [pem_pos=5, final_scores=0.46159285,
final_local_scores=0.4641549, pem=-4.481174,
similarity=42.31034]
18: 26667 Spain
    [pem_pos=1, final_scores=0.41599995,
final_local_scores=0.49362832, pem=-0.7552963,
similarity=22.927984]
19: 407139 Carlos_Sainz
    [pem_pos=0, final_scores=0.415514,
final_local_scores=0.40360135, pem=0.0, similarity=16.133982]
20: 3343 Belgium
    [pem_pos=0, final_scores=0.62491417,
final_local_scores=0.6966844, pem=-0.12403572,
similarity=25.554691]
21: 8958042 Bruno_Thiry
    [pem_pos=0, final_scores=0.62744814,
final_local_scores=0.7178692, pem=0.0, similarity=25.512362]
```

Sample: b'242_MOTOR'

```
MOTOR RACING - 1 BELGIAN GRAND PRIX GRID POSITIONS . [1 1 SP
A-FRANCORCHAMPS 1] , [2 2 Belgium 2] 1996-08-24 Grid
positions for Sunday 's 2 Belgian Grand Prix motor race after
final qualifying on Saturday : 1. [3 3 Jacques Villeneuve 3]
( [4 4 Canada 4] ) [5 5 Williams 5] 1 minute 50.574 seconds
( average speed 226.859 kph ) 2. [6 6 Damon Hill 6] ( [7 7 Br
itain 7] ) [8 8 Williams 8] 1:50.980 3. [9 9 Michael Schumach
er 9] ( [10 10 Germany 10] ) [11 11 Ferrari 11] 1:51.778 4. [1
2 12 David Coulthard 12] ( [13 13 Britain 13] ) [14 14 McLaren
```

14] 1:51.884 5. [15 15 Gerhard Berger 15] ([16 16 Austria 16]) [17 17 Benetton 17] 1:51.960 6. [18 18 Mika Hakkinen 18] ([19 19 Finland 19]) [20 20 McLaren 20] 1:52.318 7. [21 21 Jean Alesi 21] ([22 22 France 22]) [23 23 Benetton 23] 1:52.354 8. [24 24 Martin Brundle 24] ([25 25 Britain 25]) [26 26 Jordan 26] 1:52.977 9. [27 27 Eddie Irvine 27] ([28 28 Britain 28]) [29 29 Ferrari 29] 1:53.043 10. [30 30 Rubens Barrichello 30] ([31 31 Brazil 31]) [32 32 Jordan 32] 1:53.152 Add grid : 11. [33 33 Heinz-Harald Frentzen 33] ([34 34 Germany 34]) [35 35 Sauber 35] 1:53.199 12. [36 36 Johnny Herbert 36] ([37 37 Britain 37]) [38 38 Sauber 38] 1:53.993 13. [39 39 Mika Salo 39] ([40 40 Finland 40]) [41 41 Tyrrell 41] 1:54.095 14. [42 42 Olivier Panis 42] ([43 43 France 43]) [44 44 Ligier 44] 1:54.220 15. [45 45 Pedro Diniz 45] ([46 46 Brazil 46]) [47 47 Ligier 47] 1:54.700 16. [48 48 Jos Verstappen 48] ([49 49 Netherlands 49]) [50 50 Arrows 50] 1:55.150 17. [51 51 Ukyo Katayama 51] ([52 52 Japan 52]) [53 53 Tyrrell 53] 1:55.371 18. [54 54 Ricardo Rosset 54] ([55 55 Brazil 55]) [56 56 Arrows 56] 1:56.286 19. [57 57 Pedro Lamy 57] ([58 58 Portugal 58]) [59 59 Minardi 59] 1:56.830 Did not qualify (times did not meet qualifying standard) : 20. [60 60 Giovanni Lavaggi 60] ([61 61 Italy 61]) [62 62 Minardi 62] 1:58.579
1: 948477 Circuit_de_Spa-Francorchamps
2: 3343 Belgium
3: 64258 Jacques_Villeneuve
4: 5042916 Canada
5: 34104 Williams_F1
6: 78249 Damon_Hill
7: 31717 United_Kingdom
8: 34104 Williams_F1
9: 20396 Michael_Schumacher
10: 11867 Germany
11: 173009 Scuderia_Ferrari
12: 19636733 David_Coulthard
13: 31717 United_Kingdom
14: 20994 McLaren
15: 343833 Gerhard_Berger
16: 26964606 Austria
17: 445741 Benetton_Formula
18: 63073 Mika_Häkkinen
19: 10577 Finland
20: 20994 McLaren
21: 341553 Jean_Alesi
22: 5843419 France
23: 445741 Benetton_Formula
24: 341560 Martin_Brundle
25: 31717 United_Kingdom
26: 458472 Jordan_Grand_Prix
27: 424417 Eddie_Irvine
28: 31717 United_Kingdom
29: 173009 Scuderia_Ferrari

```
30: 252989 Rubens_Barrichello
31: 3383 Brazil
32: 458472 Jordan_Grand_Prix
33: 756775 Heinz-Harald_Frentzen
34: 11867 Germany
35: 3158734 Sauber
36: 342991 Johnny_Herbert
37: 31717 United_Kingdom
38: 3158734 Sauber
39: 1157079 Mika_Salo
40: 10577 Finland
41: 342015 Tyrrell_Racing
42: 414513 Olivier_Panis
43: 5843419 France
44: 24726570 Equipe_Ligier
45: 1226332 Pedro_Diniz
46: 3383 Brazil
47: 24726570 Equipe_Ligier
48: 1148735 Jos_Verstappen
49: 21148 Netherlands
50: 800001 Arrows_Grand_Prix_International
51: 520249 Ukyo_Katayama
52: 15573 Japan
53: 342015 Tyrrell_Racing
54: 1226544 Ricardo_Rosset
55: 3383 Brazil
56: 800001 Arrows_Grand_Prix_International
57: 909656 Pedro_Lamy
58: 23033 Portugal
59: 40479 Minardi
60: 1228421 Giovanni_Lavaggi
61: 14532 Italy
62: 40479 Minardi
1: 1123283 1996_Belgian_Grand_Prix
    [winner: 3343 Belgium, pem_pos=0,
final_scores=-0.13370769, final_local_scores=-0.14933945,
pem=-0.12403572, similarity=0.3076514]
    RECALL MISS: 1123283 1996_Belgian_Grand_Prix]
    [winner: 348674 Grand_Prix_motor_racing, pem_pos=0,
final_scores=-0.15614948, final_local_scores=-0.21257967,
pem=-1.5726227, similarity=6.2980785]
    RECALL MISS: 1123283 1996_Belgian_Grand_Prix]
2: 1123283 1996_Belgian_Grand_Prix
    [winner: 351087 Belgian_Grand_Prix, pem_pos=0,
final_scores=-0.052699447, final_local_scores=-0.0950905,
pem=-0.50572526, similarity=4.0022354]
    RECALL MISS: 1123283 1996_Belgian_Grand_Prix]
1: 948477 Circuit_de_Spa-Francorchamps
    [pem_pos=0, final_scores=0.3818068,
final_local_scores=0.35846618, pem=0.0, similarity=14.787059]
2: 3343 Belgium
    [pem_pos=0, final_scores=0.56733835,
```

```
final_local_scores=0.62083334, pem=-0.24180195,
similarity=23.931576]
3: 64258 Jacques_Villeneuve
    [pem_pos=0, final_scores=0.4148147,
final_local_scores=0.38729817, pem=0.0, similarity=15.647463]
4: 5042916 Canada
    [pem_pos=0, final_scores=0.67802185,
final_local_scores=0.77333045, pem=-0.11662793,
similarity=27.801676]
5: 34104 Williams_F1
    [pem_pos=0, final_scores=0.6004575,
final_local_scores=0.58208144, pem=-1.0414248,
similarity=27.123613]
6: 78249 Damon_Hill
    [pem_pos=0, final_scores=0.6008144,
final_local_scores=0.5889426, pem=0.0, similarity=21.664934]
7: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5194565,
final_local_scores=0.58497936, pem=-0.61114657,
similarity=24.870173]
8: 34104 Williams_F1
    [pem_pos=0, final_scores=0.63666904,
final_local_scores=0.62163705, pem=-1.0414248,
similarity=28.304031]
9: 20396 Michael_Schumacher
    [pem_pos=0, final_scores=0.6163596,
final_local_scores=0.598893, pem=0.0, similarity=21.961874]
10: 11867 Germany
    [pem_pos=0, final_scores=0.75352055,
final_local_scores=0.84233814, pem=-0.15529114,
similarity=30.07126]
11: 173009 Scuderia_Ferrari
    [pem_pos=1, final_scores=0.38291007,
final_local_scores=0.33855218, pem=-0.77418387,
similarity=18.402918]
12: 19636733 David_Coulthard
    [pem_pos=0, final_scores=0.7025726,
final_local_scores=0.70051694, pem=0.0, similarity=24.994535]
13: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5379596,
final_local_scores=0.6051912, pem=-0.61114657,
similarity=25.473333]
14: 20994 McLaren
    [pem_pos=0, final_scores=0.53859264,
final_local_scores=0.50607777, pem=-0.11262372,
similarity=19.804548]
15: 343833 Gerhard_Berger
    [pem_pos=0, final_scores=0.5859777,
final_local_scores=0.58069223, pem=0.0, similarity=21.418728]
16: 26964606 Austria
    [pem_pos=0, final_scores=0.61556095,
final_local_scores=0.6911811, pem=-0.25262418,
```

```
similarity=26.089746]
17: 445741 Benetton_Formula
    [pem_pos=0, final_scores=0.34929782,
final_local_scores=0.31669417, pem=-0.069357015,
similarity=13.917673]
18: 63073 Mika_Häkkinen
    [pem_pos=0, final_scores=0.80928427,
final_local_scores=0.8200288, pem=0.0, similarity=28.561007]
19: 10577 Finland
    [pem_pos=0, final_scores=0.73537785,
final_local_scores=0.83172655, pem=-0.2445894,
similarity=30.240208]
20: 20994 McLaren
    [pem_pos=0, final_scores=0.5649752,
final_local_scores=0.5348967, pem=-0.11262372,
similarity=20.664562]
21: 341553 Jean_Alesi
    [pem_pos=0, final_scores=0.53310794,
final_local_scores=0.515981, pem=0.0, similarity=19.487616]
22: 5843419 France
    [pem_pos=0, final_scores=0.6891148,
final_local_scores=0.7594153, pem=-0.16240834,
similarity=27.635382]
23: 445741 Benetton_Formula
    [pem_pos=0, final_scores=0.35873288,
final_local_scores=0.32700056, pem=-0.069357015,
similarity=14.225236]
24: 341560 Martin_Brundle
    [pem_pos=0, final_scores=0.65919304,
final_local_scores=0.67119324, pem=0.0, similarity=24.119457]
25: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5456687,
final_local_scores=0.61361223, pem=-0.61114657,
similarity=25.724634]
26: 458472 Jordan_Grand_Prix
    [pem_pos=1, final_scores=0.6771796,
final_local_scores=0.66849804, pem=-2.8876953,
similarity=39.742764]
27: 424417 Eddie_Irvine
    [pem_pos=0, final_scores=0.47420377,
final_local_scores=0.44858003, pem=0.0, similarity=17.476236]
28: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5436226,
final_local_scores=0.6113772, pem=-0.61114657,
similarity=25.657936]
29: 173009 Scuderia_Ferrari
    [pem_pos=1, final_scores=0.35925406,
final_local_scores=0.31271154, pem=-0.77418387,
similarity=17.63178]
30: 252989 Rubens_Barrichello
    [pem_pos=0, final_scores=0.52747655,
final_local_scores=0.5113163, pem=0.0, similarity=19.348412]
```

31: 3383 Brazil
[pem_pos=0, final_scores=0.65211666,
final_local_scores=0.7493834, pem=-0.21313342,
similarity=27.61186]
32: 458472 Jordan_Grand_Prix
[pem_pos=1, final_scores=0.6448931,
final_local_scores=0.63323, pem=-2.8876953,
similarity=38.690296]
33: 756775 Heinz-Harald_Frentzen
[pem_pos=0, final_scores=0.8371776,
final_local_scores=0.8534012, pem=0.0, similarity=29.556906]
34: 11867 Germany
[pem_pos=0, final_scores=0.7578295,
final_local_scores=0.847045, pem=-0.15529114,
similarity=30.211723]
35: 3158734 Sauber
[pem_pos=0, final_scores=0.38522393,
final_local_scores=0.35527778, pem=0.0, similarity=14.691911]
36: 342991 Johnny_Herbert
[pem_pos=0, final_scores=0.74515706,
final_local_scores=0.7521642, pem=0.0, similarity=26.535793]
37: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5623551,
final_local_scores=0.6318396, pem=-0.61114657,
similarity=26.268578]
38: 3158734 Sauber
[pem_pos=0, final_scores=0.37897366,
final_local_scores=0.34845033, pem=0.0, similarity=14.488166]
39: 1157079 Mika_Salo
[pem_pos=0, final_scores=0.68650705,
final_local_scores=0.7032328, pem=0.0, similarity=25.075583]
40: 10577 Finland
[pem_pos=0, final_scores=0.6898231,
final_local_scores=0.7819648, pem=-0.2445894,
similarity=28.755217]
41: 342015 Tyrrell_Racing
[pem_pos=0, final_scores=0.50006217,
final_local_scores=0.47911584, pem=-0.054408323,
similarity=18.683367]
42: 414513 Olivier_Panis
[pem_pos=0, final_scores=0.5582037,
final_local_scores=0.5654598, pem=0.0, similarity=20.96416]
43: 5843419 France
[pem_pos=0, final_scores=0.67097205,
final_local_scores=0.73959714, pem=-0.16240834,
similarity=27.043968]
44: 24726570 Equipe_Ligier
[pem_pos=1, final_scores=0.37307876,
final_local_scores=0.3475365, pem=-0.5983149,
similarity=17.714624]
45: 1226332 Pedro_Diniz
[pem_pos=0, final_scores=0.5121254,

```
final_local_scores=0.5344689, pem=0.0, similarity=20.039331]
46: 3383 Brazil
    [pem_pos=0, final_scores=0.7016973,
final_local_scores=0.80354273, pem=-0.21313342,
similarity=29.228083]
47: 24726570 Equipe_Ligier
    [pem_pos=1, final_scores=0.36910123,
final_local_scores=0.34319168, pem=-0.5983149,
similarity=17.584967]
48: 1148735 Jos_Verstappen
    [pem_pos=0, final_scores=0.48922318,
final_local_scores=0.47301626, pem=0.0, similarity=18.205462]
49: 21148 Netherlands
    [pem_pos=0, final_scores=0.56657475,
final_local_scores=0.63492113, pem=-0.24940795,
similarity=24.393345]
50: 800001 Arrows_Grand_Prix_International
    [pem_pos=9, final_scores=0.22647966,
final_local_scores=0.19186121, pem=-3.4470384,
similarity=28.560766]
51: 520249 Ukyo_Katayama
    [pem_pos=0, final_scores=0.51909417,
final_local_scores=0.5472365, pem=0.0, similarity=20.420341]
52: 15573 Japan
    [pem_pos=0, final_scores=0.70020837,
final_local_scores=0.80553406, pem=-0.13549899,
similarity=28.86532]
53: 342015 Tyrrell_Racing
    [pem_pos=0, final_scores=0.505515,
final_local_scores=0.4850722, pem=-0.054408323,
similarity=18.861116]
54: 1226544 Ricardo_Rosset
    [pem_pos=0, final_scores=0.65837866,
final_local_scores=0.7246723, pem=0.0, similarity=25.71538]
55: 3383 Brazil
    [pem_pos=0, final_scores=0.7033519,
final_local_scores=0.8053501, pem=-0.21313342,
similarity=29.282019]
56: 800001 Arrows_Grand_Prix_International
    [pem_pos=9, final_scores=0.27875012,
final_local_scores=0.2489587, pem=-3.4470384,
similarity=30.264668]
57: 909656 Pedro_Lamy
    [pem_pos=0, final_scores=0.48237735,
final_local_scores=0.501252, pem=0.0, similarity=19.048073]
58: 23033 Portugal
    [pem_pos=0, final_scores=0.58458185,
final_local_scores=0.6669423, pem=-0.30942732,
similarity=25.675314]
59: 40479 Minardi
    [pem_pos=0, final_scores=0.3164699,
final_local_scores=0.28608158, pem=0.0, similarity=12.626958]
```

```
60: 1228421 Giovanni_Lavaggi
    [pem_pos=0, final_scores=0.40194315,
final_local_scores=0.46655208, pem=0.0, similarity=18.012558]
61: 14532 Italy
    [pem_pos=0, final_scores=0.69302523,
final_local_scores=0.7752918, pem=-0.20612453,
similarity=28.346905]
62: 40479 Minardi
    [pem_pos=0, final_scores=0.3387698,
final_local_scores=0.3104408, pem=0.0, similarity=13.353886]
```

Sample: b'243_RALLYING'

RALLYING - [1 1 BELGIAN 1] SPECTATOR DIES IN [2 2 FINNISH 2]
RALLY . [3 3 JYVASKYLA 3] , [4 4 Finland 4] 1996-08-24 A [5 5
Belgian 5] man died and 31 people were injured after an
accident in Friday 's opening phase of the world championship
1,000 Lakes Rally . The unnamed victim died during the
night , a hospital spokesman said on Saturday . [6 6 Danish 6]
driver Karsten Richardt had ploughed into the crowd during
the two-kilometre first stage held in the host city of [7 7 Jy
vaskyla 7] . Richardt 's [8 8 Mitsubishi 8] skidded down an
escape road and ploughed into a cordoned-off area for
spectators . A second [9 9 Belgian 9] was also seriously
injured but the hospital spokesman said his life was not in
danger . The stage was suspended but the four-day rally
resumed on Saturday . A woman was killed before last year 's
event when she walked in front of a car practising on the
course .

```
1: 3343 Belgium
2: 10577 Finland
3: 258746 Jyväskylä
4: 10577 Finland
5: 3343 Belgium
6: 76972 Denmark
7: 258746 Jyväskylä
8: 203872 Mitsubishi_Motors
9: 3343 Belgium
1: 3343 Belgium
    [pem_pos=0, final_scores=0.5291982,
final_local_scores=0.49569303, pem=-0.12403572,
similarity=19.556707]
2: 10577 Finland
    [pem_pos=0, final_scores=0.46910316,
final_local_scores=0.41758364, pem=-0.8112885,
similarity=20.963154]
3: 258746 Jyväskylä
    [pem_pos=0, final_scores=0.55893975,
final_local_scores=0.57771826, pem=0.0, similarity=21.329979]
4: 10577 Finland
    [pem_pos=0, final_scores=0.8452516,
final_local_scores=0.8284688, pem=-0.2445894,
similarity=30.14299]
```

```
5: 3343 Belgium
    [pem_pos=0, final_scores=0.7262911,
final_local_scores=0.71098703, pem=-0.12403572,
similarity=25.981508]
6: 76972 Denmark
    [pem_pos=0, final_scores=0.62513846,
final_local_scores=0.6049842, pem=-0.65021485,
similarity=25.679619]
7: 258746 Jyväskylä
    [pem_pos=0, final_scores=0.39516175,
final_local_scores=0.39881563, pem=0.0, similarity=15.991166]
8: 203872 Mitsubishi_Motors
    [pem_pos=1, final_scores=0.22248666,
final_local_scores=0.34036648, pem=-1.328055,
similarity=21.469097]
9: 3343 Belgium
    [pem_pos=0, final_scores=0.6428233,
final_local_scores=0.6198112, pem=-0.12403572,
similarity=23.26064]
```

Sample: b'244_TENNIS'

```
TENNIS - RESULTS AT [1 1 TOSHIBA CLASSIC 1] . [2 2 CARLSBAD 2
] , [3 3 California 3] 1996-08-24 Results from the $ 450,000 [
4 4 Toshiba Classic 4] tennis tournament on Friday ( prefix
number denotes seeding ) : Quarterfinals : 1 - [5 5 Arantxa Sa
nchez Vicario 5] ( [6 6 Spain 6] ) beat [7 7 Katarina Studeni
kova 7] ( [8 8 Slovakia 8] ) 6-3 6-3 3 - [9 9 Jana Novotna 9]
( [10 10 Czech Republic 10] ) beat [11 11 Sandrine Testud 11]
( [12 12 France 12] ) 2-6 7-6 ( 7-4 ) 6-3 . 4 - [13 13 Kimiko
Date 13] ( [14 14 Japan 14] ) beat 5 - [15 15 Gabriela Sabati
ni 15] ( [16 16 Argentina 16] ) 6-4 6-1
1: 31726664 Southern_California_Open
2: 108030 Carlsbad,_California
3: 5407 California
4: 31726664 Southern_California_Open
5: 534624 Arantxa_Sánchez_Vicario
6: 26667 Spain
7: 15290332 Katarína_Studeníková
8: 26830 Slovakia
9: 307050 Jana_Novotná
10: 5321 Czech_Republic
11: 3127711 Sandrine_Testud
12: 5843419 France
13: 495296 Kimiko_Date-Krumm
14: 15573 Japan
15: 782187 Gabriela_Sabatini
16: 18951905 Argentina
1: 31726664 Southern_California_Open
    [pem_pos=0, final_scores=0.4547491,
final_local_scores=0.49138218, pem=0.0, similarity=18.753538]
2: 108030 Carlsbad,_California
    [pem_pos=0, final_scores=0.50373447,
```

```
final_local_scores=0.5650581, pem=-0.36353475,
similarity=22.929132]
3: 5407 California
    [pem_pos=0, final_scores=0.6116268,
final_local_scores=0.69634396, pem=-0.08662026,
similarity=25.34106]
4: 31726664 Southern_California_Open
    [pem_pos=0, final_scores=0.54135364,
final_local_scores=0.5859845, pem=0.0, similarity=21.57666]
5: 534624 Arantxa_Sánchez_Vicario
    [pem_pos=0, final_scores=0.57901245,
final_local_scores=0.62627923, pem=0.0, similarity=22.779135]
6: 26667 Spain
    [pem_pos=0, final_scores=0.58361703,
final_local_scores=0.618832, pem=-0.18857051,
similarity=23.582369]
7: 15290332 Katarína_Studeníková
    [pem_pos=0, final_scores=0.5966782,
final_local_scores=0.7552054, pem=0.0, similarity=26.626549]
8: 26830 Slovakia
    [pem_pos=0, final_scores=0.5830934,
final_local_scores=0.6150485, pem=-0.36602598,
similarity=24.434494]
9: 307050 Jana_Novotná
    [pem_pos=0, final_scores=0.88556355,
final_local_scores=0.97172755, pem=0.0, similarity=33.088]
10: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.5476211,
final_local_scores=0.57954746, pem=-0.27794555,
similarity=22.896076]
11: 3127711 Sandrine_Testud
    [pem_pos=0, final_scores=0.7201516,
final_local_scores=0.81378657, pem=0.0, similarity=28.374727]
12: 5843419 France
    [pem_pos=0, final_scores=0.7313354,
final_local_scores=0.78261477, pem=-0.16240834,
similarity=28.3277]
13: 495296 Kimiko_Date-Krumm
    [pem_pos=0, final_scores=0.70829594,
final_local_scores=0.77541614, pem=0.0, similarity=27.229677]
14: 15573 Japan
    [pem_pos=0, final_scores=0.65653664,
final_local_scores=0.72839314, pem=-0.13549899,
similarity=26.56328]
15: 782187 Gabriela_Sabatini
    [pem_pos=0, final_scores=0.83678323,
final_local_scores=0.9048285, pem=0.0, similarity=31.091599]
16: 18951905 Argentina
    [pem_pos=0, final_scores=0.63589954,
final_local_scores=0.6734274, pem=-0.30003807,
similarity=25.817783]
```

Sample: b'245_TENNIS'

TENNIS - RESULTS AT HAMLET CUP . [1 1 COMMACK 1] , [2 2 New York 2] 1996-08-23 Results from the Waldbaum Hamlet Cup tennis tournament on Friday (prefix number denotes seeding) :
Quarterfinals : [3 3 Adrian Voinea 3] ([4 4 Romania 4]) beat [5 5 Thomas Johansson 5] ([6 6 Sweden 6]) 7-6 (7-4) 6-2 5 - [7 7 Andrey Medvedev 7] ([8 8 Ukraine 8]) beat [9 9 Jonathan Stark 9] ([10 10 U.S. 10]) 7-6 (7-4) 4-6 6-3 [11 11 Martin Damm 11] ([12 12 Czech Republic 12]) beat [13 13 Michael Joyce 13] ([14 14 U.S. 14]) 5-7 6-2 6-3 [15 15 Karol Kucera 15] ([16 16 Slovakia 16]) beat 1 - [17 17 Michael Chang 17] ([18 18 U.S. 18]) 6-4 6-4

1: 127266 Commack,_New_York
2: 8210131 New_York
3: 4036370 Adrian_Voinea
4: 25445 Romania
5: 1408278 Thomas_Johansson
6: 5058739 Sweden
7: 702849 Andriy_Medvedev
8: 31750 Ukraine
9: 3079144 Jonathan_Stark_(tennis)
10: 3434750 United_States
11: 4026614 Martin_Damm
12: 5321 Czech_Republic
13: 4026170 Michael_Joyce_(tennis)
14: 3434750 United_States
15: 3039956 Karol_Kučera
16: 26830 Slovakia
17: 391834 Michael_Chang
18: 3434750 United_States
1: 127266 Commack,_New_York
[pem_pos=0, final_scores=0.49938834, final_local_scores=0.6292941, pem=0.0, similarity=22.869104]
2: 8210131 New_York
[pem_pos=0, final_scores=0.17661242, final_local_scores=0.17766187, pem=-0.2889045, similarity=10.9626045]
3: 4036370 Adrian_Voinea
[pem_pos=0, final_scores=0.5465546, final_local_scores=0.67560375, pem=0.0, similarity=24.251076]
4: 25445 Romania
[pem_pos=0, final_scores=0.69656616, final_local_scores=0.72713333, pem=-0.21241479, similarity=26.943966]
5: 1408278 Thomas_Johansson
[pem_pos=0, final_scores=0.8173818, final_local_scores=0.91549593, pem=0.0, similarity=31.409937]
6: 5058739 Sweden
[pem_pos=0, final_scores=0.7497748, final_local_scores=0.77750325, pem=-0.2135508, similarity=28.453283]
7: 702849 Andriy_Medvedev

```
[pem_pos=0, final_scores=0.97020656,
final_local_scores=1.085973, pem=0.0, similarity=36.49731]
8: 31750 Ukraine
[pem_pos=0, final_scores=0.61851674,
final_local_scores=0.65034825, pem=-0.35407752,
similarity=25.42293]
9: 3079144 Jonathan_Stark_(tennis)
[pem_pos=1, final_scores=0.59894145,
final_local_scores=0.7063499, pem=-0.86421156,
similarity=29.868319]
10: 3434750 United_States
[pem_pos=0, final_scores=0.52764034,
final_local_scores=0.52618927, pem=-0.14075686,
similarity=20.557707]
11: 4026614 Martin_Damm
[pem_pos=0, final_scores=0.89946395,
final_local_scores=1.024641, pem=0.0, similarity=34.667046]
12: 5321 Czech_Republic
[pem_pos=0, final_scores=0.5760092,
final_local_scores=0.58012426, pem=-0.27794555,
similarity=22.91329]
13: 4026170 Michael_Joyce_(tennis)
[pem_pos=1, final_scores=0.49063814,
final_local_scores=0.58883554, pem=-0.41227743,
similarity=23.903769]
14: 3434750 United_States
[pem_pos=0, final_scores=0.5231148,
final_local_scores=0.5212458, pem=-0.14075686,
similarity=20.410183]
15: 3039956 Karol_Kučera
[pem_pos=0, final_scores=0.7936041,
final_local_scores=0.9042041, pem=0.0, similarity=31.072966]
16: 26830 Slovakia
[pem_pos=0, final_scores=0.7003251,
final_local_scores=0.7158246, pem=-0.36602598,
similarity=27.441854]
17: 391834 Michael_Chang
[pem_pos=0, final_scores=0.7569466,
final_local_scores=0.8216166, pem=0.0, similarity=28.60839]
18: 3434750 United_States
[pem_pos=0, final_scores=0.5382814,
final_local_scores=0.537813, pem=-0.14075686,
similarity=20.904583]
```

Sample: b'246_GOLF'
GOLF - THUNDERSTORMS FORCE SUSPENSION OF SECOND ROUND IN [1 1
AKRON 1] . [2 2 AKRON 2] , [3 3 Ohio 3] 1996-08-23
Thunderstorms forced the suspension of the [4 4 World Series o
f Golf 4] after just five players in the 43-man field had
completed the second round on Friday . Play initially was
interrupted for more than 3-1/2 hours before resuming for two
hours . But play was finally suspended for the day when the

storm continued . The 38 players are scheduled to resume their rounds on Saturday morning before the third round begins .

```
1: 129842 Akron,_Ohio
2: 129842 Akron,_Ohio
3: 22199 Ohio
4: 1536678 World_Series_of_Golf
1: 129842 Akron,_Ohio
    [pem_pos=0, final_scores=0.46379,
final_local_scores=0.41431153, pem=-0.6020413,
similarity=19.72759]
2: 129842 Akron,_Ohio
    [pem_pos=0, final_scores=0.67978543,
final_local_scores=0.6502537, pem=-0.6020413,
similarity=26.768576]
3: 22199 Ohio
    [pem_pos=0, final_scores=0.65816945,
final_local_scores=0.6696166, pem=-0.19147073,
similarity=25.113655]
4: 1536678 World_Series_of_Golf
    [pem_pos=0, final_scores=0.31231958,
final_local_scores=0.43772793, pem=0.0, similarity=17.152388]
```

Sample: b'247_GOLF'

GOLF - SCORES AT [1 1 WORLD SERIES OF GOLF 1] . [2 2 AKRON 2]
, [3 3 Ohio 3] 1996-08-23 Scores from the \$ 2.1 million [4 4 N
EC World Series of Golf 4] after the second round was
suspended due to rain on Friday with 38 of the 43 players yet
to finish their rounds on the 7,149 yard , par 70 Firestone
C.C course (players [5 5 U.S. 5] unless noted) : - 5 [6 6 Pa
ul Goydos 6] through 2 holes - 5 [7 7 Billy Mayfair 7]
through 2 - 4 [8 8 Greg Norman 8] ([9 9 Australia 9])
through 8 - 4 [10 10 Hidemichi Tanaka 10] ([11 11 Japan
11]) through 3 - 3 [12 12 Steve Stricker 12] through 3 - 2 [1
3 13 Phil Mickelson 13] through 7 - 2 [14 14 Mark Brooks 14]
through 3 - 1 [15 15 Hal Sutton 15] through 11 - 1 [16 16 John
Cook 16] through 5 - 1 [17 17 Tim Herron 17] through 4 - 1 [18
18 Justin Leonard 18] through 3 0 [19 19 Steve Jones 19]
through 7 0 [20 20 Nick Faldo 20] ([21 21 Britain 21])
through 5 0 [22 22 Davis Love 22] through 5 +1 [23 23 Fred Co
uples 23] through 15 +1 [24 24 Fred Funk 24] through 9 +1 [25
25 Scott Hoch 25] through 9 +1 [26 26 Ernie Els 26] ([27 27 S
outh Africa 27]) through 8 +2 [28 28 D.A. Weibring 28]
through 12 +2 [29 29 Clarence Rose 29] through 9 +2 [30 30 Du
ffy Waldorf 30] through 4 +3 [31 31 Jim Furyk 31] through 16
+3 [32 32 Corey Pavin 32] through 14 +3 [33 33 Craig Stadler 3
3] through 14 +3 [34 34 Brad Bryant 34] through 12 +3 [35 35 T
om Lehman 35] through 11 +3 [36 36 Sven Struver 36] ([37 37 G
ermany 37]) through 10 +3 [38 38 Alexander Cejka 38] ([39 39
Germany 39]) through 10 +3 [40 40 Anders Forsbrand 40] ([41
41 Sweden 41]) through 5 +4 [42 42 Willie Wood 42] through
17 +4 [43 43 Costantino Rocca 43] ([44 44 Italy 44])

through 15 +4 [45 45 Stewart Ginn 45] ([46 46 Australia
46]) through 13 +5 [47 47 Wayne Westner 47] ([48 48 South Af
rica 48]) 77 68 +5 Sigeki Maruyama ([49 49 Japan 49])
through 17 +5 Mark O'Meara through 15 +5 [50 50 Loren Roberts
50] through 9 +6 [51 51 Scott McCarron 51] 76 70 +7 [52 52 Sa
toshi Higashi 52] ([53 53 Japan 53]) through 16 +7 [54 54 Pa
ul Stankowski 54] through 15 +8 [55 55 Craig Parry 55] ([56 5
6 Australia 56]) through 13 +9 [57 57 Tom Watson 57] 79 70
+11 Seiki Okuda ([58 58 Japan 58]) 81 70 (through 18) +11
Steve Schneiter 77 74) through 18)
1: 1536678 World_Series_of_Golf
2: 129842 Akron,_Ohio
3: 22199 Ohio
4: 1527695 WGC-Bridgestone_Invitational
5: 3434750 United_States
6: 6280981 Paul_Goydos
7: 2080225 Billy_Mayfair
8: 157447 Greg_Norman
9: 4689264 Australia
10: 2114665 Hidemichi_Tanaka
11: 15573 Japan
12: 2310551 Steve_Stricker
13: 592392 Phil_Mickelson
14: 2295891 Mark_Brooks_(golfer)
15: 654136 Hal_Sutton
16: 2810777 John_Cook_(golfer)
17: 5242796 Tim_Herron
18: 1507170 Justin_Leonard
19: 1924538 Steve_Jones_(golfer)
20: 759285 Nick_Faldo
21: 31717 United_Kingdom
22: 1388818 Davis_Love_III
23: 1221065 Fred_Couples
24: 1646050 Fred_Funk
25: 3396339 Scott_Hoch
26: 230467 Ernie_Els
27: 17416221 South_Africa
28: 5653062 D._A._Weibring
29: 8385080 Clarence_Rose
30: 1781343 Duffy_Waldorf
31: 247955 Jim_Furyk
32: 1530024 Corey_Pavin
33: 1261748 Craig_Stadler
34: 4450518 Brad_Bryant
35: 1516564 Tom_Lehman
36: 4337089 Sven_Strüver
37: 11867 Germany
38: 2080183 Alex_Čejka
39: 11867 Germany
40: 4377636 Anders_Forsbrand
41: 5058739 Sweden
42: 7018258 Willie_Wood_(golfer)

```
43: 2157151 Costantino_Rocca
44: 14532 Italy
45: 5418317 Stewart_Ginn
46: 4689264 Australia
47: 5206729 Wayne_Westner
48: 17416221 South_Africa
49: 15573 Japan
50: 2567564 Loren_Roberts
51: 5845032 Scott_McCarron
52: 21138421 Satoshi_Higashi
53: 15573 Japan
54: 6189303 Paul_Stankowski
55: 3738086 Craig_Parry
56: 4689264 Australia
57: 585794 Tom_Watson_(golfer)
58: 15573 Japan
1: 1536678 World_Series_of_Golf
    [pem_pos=0, final_scores=0.42029262,
final_local_scores=0.48783576, pem=0.0, similarity=18.647705]
2: 129842 Akron,_Ohio
    [pem_pos=0, final_scores=0.56383795,
final_local_scores=0.63121784, pem=-0.6020413,
similarity=26.200506]
3: 22199 Ohio
    [pem_pos=0, final_scores=0.5833604,
final_local_scores=0.6722222, pem=-0.19147073,
similarity=25.191412]
4: 1527695 WGC-Bridgestone_Invitational
    [pem_pos=0, final_scores=0.39776158,
final_local_scores=0.4139458, pem=0.0, similarity=16.44268]
5: 3434750 United_States
    [pem_pos=0, final_scores=0.44182402,
final_local_scores=0.4939645, pem=-0.14075686,
similarity=19.596058]
6: 6280981 Paul_Goydos
    [pem_pos=0, final_scores=0.73133117,
final_local_scores=0.7785377, pem=0.0, similarity=27.32283]
7: 2080225 Billy_Mayfair
    [pem_pos=0, final_scores=0.73464864,
final_local_scores=0.78154534, pem=0.0, similarity=27.412584]
8: 157447 Greg_Norman
    [pem_pos=0, final_scores=0.6407835,
final_local_scores=0.6532412, pem=0.0, similarity=23.583735]
9: 4689264 Australia
    [pem_pos=0, final_scores=0.6378409,
final_local_scores=0.7132166, pem=-0.17168477,
similarity=26.307169]
10: 2114665 Hidemichi_Tanaka
    [pem_pos=0, final_scores=0.48986602,
final_local_scores=0.6268245, pem=0.0, similarity=22.795406]
11: 15573 Japan
    [pem_pos=0, final_scores=0.68492836,
```

```
final_local_scores=0.77116495, pem=-0.13549899,  
similarity=27.839678]  
12: 2310551 Steve_Stricker  
    [pem_pos=0, final_scores=0.7379768,  
final_local_scores=0.77212924, pem=0.0, similarity=27.13159]  
13: 592392 Phil_Mickelson  
    [pem_pos=0, final_scores=0.8339192,  
final_local_scores=0.8638565, pem=0.0, similarity=29.868914]  
14: 2295891 Mark_Brooks_(golfer)  
    [pem_pos=0, final_scores=0.638238,  
final_local_scores=0.6825135, pem=-0.17831625,  
similarity=25.426987]  
15: 654136 Hal_Sutton  
    [pem_pos=0, final_scores=0.5543614,  
final_local_scores=0.5803838, pem=0.0, similarity=21.409523]  
16: 2810777 John_Cook_(golfer)  
    [pem_pos=1, final_scores=0.5210539,  
final_local_scores=0.52807915, pem=-1.1074443,  
similarity=25.871101]  
17: 5242796 Tim_Herron  
    [pem_pos=0, final_scores=0.7544102,  
final_local_scores=0.8116688, pem=0.0, similarity=28.31153]  
18: 1507170 Justin_Leonard  
    [pem_pos=0, final_scores=0.84671015,  
final_local_scores=0.8922728, pem=0.0, similarity=30.716913]  
19: 1924538 Steve_Jones_(golfer)  
    [pem_pos=4, final_scores=0.43845564,  
final_local_scores=0.46138757, pem=-2.187635,  
similarity=29.755136]  
20: 759285 Nick_Faldo  
    [pem_pos=0, final_scores=0.8661944,  
final_local_scores=0.89743954, pem=0.0, similarity=30.8711]  
21: 31717 United_Kingdom  
    [pem_pos=0, final_scores=0.53473383,  
final_local_scores=0.6182669, pem=-0.61114657,  
similarity=25.863539]  
22: 1388818 Davis_Love_III  
    [pem_pos=0, final_scores=0.85800225,  
final_local_scores=0.89070636, pem=0.0, similarity=30.670168]  
23: 1221065 Fred_Couples  
    [pem_pos=0, final_scores=0.6376761,  
final_local_scores=0.65718085, pem=0.0, similarity=23.7013]  
24: 1646050 Fred_Funk  
    [pem_pos=0, final_scores=0.5814176,  
final_local_scores=0.60775614, pem=0.0, similarity=22.226368]  
25: 3396339 Scott_Hoch  
    [pem_pos=0, final_scores=0.7387646,  
final_local_scores=0.7925648, pem=0.0, similarity=27.741428]  
26: 230467 Ernie_Els  
    [pem_pos=0, final_scores=0.73690933,  
final_local_scores=0.76804566, pem=0.0, similarity=27.009727]  
27: 17416221 South_Africa
```

```
[pem_pos=0, final_scores=0.5296586,
final_local_scores=0.61115897, pem=-0.26251382,
similarity=23.755507]
28: 5653062 D._A._Weibring
    [pem_pos=0, final_scores=0.514111,
final_local_scores=0.5605926, pem=0.0, similarity=20.818914]
29: 8385080 Clarence_Rose
    [pem_pos=0, final_scores=0.59982955,
final_local_scores=0.6947939, pem=0.0, similarity=24.823748]
30: 1781343 Duffy_Waldorf
    [pem_pos=0, final_scores=0.6098611,
final_local_scores=0.68583524, pem=0.0, similarity=24.556404]
31: 247955 Jim_Furyk
    [pem_pos=0, final_scores=0.8836619,
final_local_scores=0.91958755, pem=0.0, similarity=31.53204]
32: 1530024 Corey_Pavin
    [pem_pos=0, final_scores=0.8879593,
final_local_scores=0.92215747, pem=0.0, similarity=31.60873]
33: 1261748 Craig_Stadler
    [pem_pos=0, final_scores=0.7348749,
final_local_scores=0.771682, pem=0.0, similarity=27.118244]
34: 4450518 Brad_Bryant
    [pem_pos=0, final_scores=0.6827369,
final_local_scores=0.73165494, pem=0.0, similarity=25.923756]
35: 1516564 Tom_Lehman
    [pem_pos=0, final_scores=0.8901965,
final_local_scores=0.9233833, pem=0.0, similarity=31.645311]
36: 4337089 Sven_Strüver
    [pem_pos=0, final_scores=0.6907623,
final_local_scores=0.79086226, pem=0.0, similarity=27.69062]
37: 11867 Germany
    [pem_pos=0, final_scores=0.7713199,
final_local_scores=0.8728011, pem=-0.15529114,
similarity=30.980337]
38: 2080183 Alex_Čejka
    [pem_pos=0, final_scores=0.7682467,
final_local_scores=0.82440525, pem=0.0, similarity=28.69161]
39: 11867 Germany
    [pem_pos=0, final_scores=0.77459556,
final_local_scores=0.8763793, pem=-0.15529114,
similarity=31.087118]
40: 4377636 Anders_Forsbrand
    [pem_pos=0, final_scores=0.5668287,
final_local_scores=0.6245279, pem=0.0, similarity=22.72687]
41: 5058739 Sweden
    [pem_pos=0, final_scores=0.6557159,
final_local_scores=0.7343775, pem=-0.2135508,
similarity=27.166325]
42: 7018258 Willie_Wood_(golfer)
    [pem_pos=1, final_scores=0.56616837,
final_local_scores=0.6112638, pem=-0.86820155,
similarity=27.05246]
```

43: 2157151 Costantino_Rocca
[pem_pos=0, final_scores=0.5898969,
final_local_scores=0.6282605, pem=0.0, similarity=22.838259]
44: 14532 Italy
[pem_pos=0, final_scores=0.6332659,
final_local_scores=0.7438829, pem=-0.20612453,
similarity=27.4096]
45: 5418317 Stewart_Ginn
[pem_pos=0, final_scores=0.42962533,
final_local_scores=0.5256368, pem=0.0, similarity=19.775763]
46: 4689264 Australia
[pem_pos=0, final_scores=0.64342564,
final_local_scores=0.719317, pem=-0.17168477,
similarity=26.489216]
47: 5206729 Wayne_Westner
[pem_pos=0, final_scores=0.71667933,
final_local_scores=0.80263805, pem=0.0, similarity=28.042034]
48: 17416221 South_Africa
[pem_pos=0, final_scores=0.49025267,
final_local_scores=0.568114, pem=-0.26251382,
similarity=22.470959]
49: 15573 Japan
[pem_pos=0, final_scores=0.69031763,
final_local_scores=0.7770519, pem=-0.13549899,
similarity=28.015356]
50: 2567564 Loren_Roberts
[pem_pos=0, final_scores=0.6104731,
final_local_scores=0.6333755, pem=0.0, similarity=22.990902]
51: 5845032 Scott_McCarron
[pem_pos=0, final_scores=0.674026,
final_local_scores=0.7396246, pem=0.0, similarity=26.161587]
52: 21138421 Satoshi_Higashi
[pem_pos=0, final_scores=0.51997375,
final_local_scores=0.63024235, pem=0.0, similarity=22.897402]
53: 15573 Japan
[pem_pos=0, final_scores=0.70693874,
final_local_scores=0.795208, pem=-0.13549899,
similarity=28.557169]
54: 6189303 Paul_Stankowski
[pem_pos=0, final_scores=0.68164253,
final_local_scores=0.76693696, pem=0.0, similarity=26.97664]
55: 3738086 Craig_Parry
[pem_pos=0, final_scores=0.74605125,
final_local_scores=0.7819816, pem=0.0, similarity=27.425604]
56: 4689264 Australia
[pem_pos=0, final_scores=0.6650228,
final_local_scores=0.74290866, pem=-0.17168477,
similarity=27.193237]
57: 585794 Tom_Watson_(golfer)
[pem_pos=0, final_scores=0.65675837,
final_local_scores=0.67491686, pem=-0.48845047,
similarity=26.886848]

```

58: 15573 Japan
    [pem_pos=0, final_scores=0.7198831,
final_local_scores=0.8093477, pem=-0.13549899,
similarity=28.979126]

Sample: b'248_RUGBY'
[1 1 RUGBY LEAGUE 1] - EUROPEAN SUPER LEAGUE RESULTS /
STANDINGS . [2 2 LONDON 2] 1996-08-24 Results of [3 3 European
Super League 3] rugby league matches on Saturday : 1 Paris 14
[4 4 Bradford 4] 27 [5 5 Wigan 5] 78 [6 6 Workington 6] 4
Standings ( tabulated under played , won , drawn , lost ,
points for , against , total points ) : [7 7 Wigan 7] 22 19 1
2 902 326 39 [8 8 St Helens 8] 21 19 0 2 884 441 38 [9 9 Brad
ford 9] 22 17 0 5 767 409 34 [10 10 Warrington 10] 21 12 0 9
555 499 24 2 London 21 11 1 9 555 462 23 3 Sheffield 21 10 0
11 574 696 20 [11 11 Halifax 11] 21 9 1 11 603 552 19 [12 12 C
astleford 12] 21 9 0 12 548 543 18 [13 13 Oldham 13] 21 8 1
12 439 656 17 [14 14 Leeds 14] 21 6 0 15 531 681 12 4 Paris
22 3 1 18 398 795 7 [15 15 Workington 15] 22 2 1 19 325 1021 5
1: 25735 Rugby_league
2: 17867 London
3: 22751937 Super_League
4: 1095011 Bradford_Bulls
5: 1095029 Wigan_Warriors
6: 3646900 Workington_Town
7: 1095029 Wigan_Warriors
8: 1095032 St_Helens_RLFC
9: 1095011 Bradford_Bulls
10: 852392 Warrington_Wolves
11: 1196374 Halifax_RLFC
12: 1216025 Castleford_Tigers
13: 1098719 Oldham_Roughyeds
14: 1095024 Leeds_Rhinos
15: 3646900 Workington_Town
1: 5690457 Paris_Saint-Germain_Rugby_League
    [winner: 22989 Paris, pem_pos=0,
final_scores=-0.52904594, final_local_scores=-0.4954659,
pem=-0.115301244, similarity=-10.068952]
    RECALL MISS: 5690457 Paris_Saint-Germain_Rugby_League]
2: 629563 London_Wasps
    [winner: 17867 London, pem_pos=0,
final_scores=-0.51509446, final_local_scores=-0.5239282,
pem=-0.09086737, similarity=-11.051199]
    RECALL MISS: 629563 London_Wasps]
3: 3367769 Sheffield_Eagles
    [winner: 88885 Sheffield, pem_pos=0,
final_scores=-0.46198368, final_local_scores=-0.49485952,
pem=-0.25810194, similarity=-9.274284]
    RECALL MISS: 3367769 Sheffield_Eagles]
4: 5690457 Paris_Saint-Germain_Rugby_League
    [winner: 22989 Paris, pem_pos=0,
final_scores=-0.50909823, final_local_scores=-0.47367606,

```

pem=-0.115301244, similarity=-9.418699 |
RECALL MISS: 5690457 Paris_Saint-Germain_Rugby_League]

1: 25735 Rugby_league
[pem_pos=0, final_scores=0.3176344,
final_local_scores=0.29765907, pem=0.0, similarity=12.972453]

2: 17867 London
[pem_pos=0, final_scores=0.54615337,
final_local_scores=0.6558767, pem=-0.09086737,
similarity=24.156532]

3: 22751937 Super_League
[pem_pos=0, final_scores=0.47283894,
final_local_scores=0.4294693, pem=0.0, similarity=16.905933]

4: 1095011 Bradford_Bulls
[pem_pos=9, final_scores=0.7708803,
final_local_scores=0.7309337, pem=-3.818427,
similarity=46.667435]

5: 1095029 Wigan_Warriors
[pem_pos=1, final_scores=0.52265316,
final_local_scores=0.46496123, pem=-1.8027086,
similarity=27.768494]

6: 3646900 Workington_Town
[pem_pos=3, final_scores=0.4468035,
final_local_scores=0.39110827, pem=-1.9682475,
similarity=26.4648]

7: 1095029 Wigan_Warriors
[pem_pos=1, final_scores=0.5593179,
final_local_scores=0.50501186, pem=-1.8027086,
similarity=28.963684]

8: 1095032 St_Helens_RLFC
[pem_pos=0, final_scores=0.71721095,
final_local_scores=0.67876023, pem=-1.1285291,
similarity=30.48239]

9: 1095011 Bradford_Bulls
[pem_pos=9, final_scores=0.7053869,
final_local_scores=0.6593922, pem=-3.818427,
similarity=44.53249]

10: 852392 Warrington_Wolves
[pem_pos=1, final_scores=0.80606824,
final_local_scores=0.772868, pem=-1.6982093,
similarity=36.38876]

11: 1196374 Halifax_RLFC
[pem_pos=3, final_scores=0.7895089,
final_local_scores=0.759113, pem=-2.7129183,
similarity=41.49643]

12: 1216025 Castleford_Tigers
[pem_pos=1, final_scores=0.7136896,
final_local_scores=0.681925, pem=-0.78848183,
similarity=28.727602]

13: 1098719 Oldham_Roughyeds
[pem_pos=1, final_scores=0.59008086,
final_local_scores=0.5469318, pem=-2.0696447,
similarity=31.666296]

```
14: 1095024 Leeds_Rhinos
    [pem_pos=1, final_scores=0.57619566,
final_local_scores=0.5213235, pem=-2.7286434,
similarity=34.485836]
15: 3646900 Workington_Town
    [pem_pos=3, final_scores=0.44693416,
final_local_scores=0.39125097, pem=-1.9682475,
similarity=26.469059]
```

Sample: b'249_CRICKET'

CRICKET - [1 1 ESSEX 1] POISED TO STEP UP TITLE CHALLENGE . [2 2 LONDON 2] 1996-08-24 [3 3 Essex 3] are set to step up their [4 4 English 4] county championship challenge with a fifth consecutive victory after new-ball pair [5 5 Neil Williams 5] and [6 6 Mark Ilott 6] sent [7 7 Gloucestershire 7] reeling on Saturday . Williams seized two wickets in two deliveries and left-armer [8 8 Ilott 8] also captured two as [9 9 Gloucestershire 9] , 252 behind on first innings , slumped to 27 for four at the close on the third day of the four-day game at [10 10 Colchester 10] . [11 11 Essex 11] , who started the current round of matches in fifth place 20 points behind leaders [12 12 Derbyshire 12] with a game in hand , could go top if they complete victory on Monday 's last day and other results favour them . Williams thrust [13 1 3 Essex 13] on course for success by dispatching [14 14 Matt Windous 14] and [15 15 Andrew Symonds 15] in his second over , before [16 16 Ilott 16] removed [17 17 Dominic Hewson 17] and [18 18 Tim Hancock 18] to reduce [19 19 Gloucestershire 19] to 17 for four at one stage . The visitors had started the day optimistically by sending back former [20 20 England 20] captain [21 21 Graham Gooch 21] when he added just six to his overnight 105 , but [22 22 Essex 22] went on to make 532 for eight before declaring . Captain [23 23 Paul Prichard 23] plundered 88 from 73 deliveries , hitting 15 fours and one six . Second-placed [24 24 Kent 24] were frustrated by rain which prevented any play at [25 25 Cardiff 25] , where they have reached 255 for three in their first innings against [26 26 Glamorgan 26] , while third-placed [27 27 Surrey 27] are facing an uphill task against [28 28 Nottinghamshire 28] . [29 29 Surrey 29] slipped to 88 for four in reply to [30 30 Notts 30] ' commanding first innings of 446 for nine declared at [31 31 Trent Bridge 31] , before [32 32 Alistair Brown 32] struck a 55-ball half-century . [33 33 Brown 33] 's 56 not out , containing eight fours and three sixes , lifted [34 34 Surrey 34] to 128 for four on a rain-curtailed day . Fourth-placed [35 35 Leicestershire 35] had [36 36 Hampshire 36] on the ropes at [37 37 Leicester 37] before rain intervened . Pace trio [38 38 David Millns 38] (2-24) , [39 39 Gordon Parsons 39] (3-20) and [40 40 Vince Wells 40] (2-19) had [41 41 Hampshire 41] reeling at 81 for seven in reply to the home county 's first innings of 353 . [42 42 Yorkshire 42] rekindled their title hopes after three successive defeats by taking

the upper hand against arch-rivals [43 43 Lancashire 43] at [4 4 44 Old Trafford 44]. Facing [45 45 Yorkshire 45]'s 529 for eight declared, [46 46 Lancashire 46] were forced to follow on 206 behind after being bowled out for 323, paceman [47 47 Darren Gough 47] polishing off the innings with a burst of three wickets for one run in 17 deliveries. [48 48 Lancashire 48] then reached 210 for five at the close -- just four ahead -- after [49 49 Neil Fairbrother 49] hit 55 in 60 balls to add to his first innings of 86.

- 1: 1622252 Essex_County_Cricket_Club
- 2: 17867 London
- 3: 1622252 Essex_County_Cricket_Club
- 4: 9316 England
- 5: 2856979 Neil_Williams_(cricketer)
- 6: 3445394 Mark_Ilott
- 7: 1622322 Gloucestershire_County_Cricket_Club
- 8: 3445394 Mark_Ilott
- 9: 1622322 Gloucestershire_County_Cricket_Club
- 10: 264105 Colchester
- 11: 1622252 Essex_County_Cricket_Club
- 12: 1829984 Derbyshire_County_Cricket_Club
- 13: 1622252 Essex_County_Cricket_Club
- 14: 6136803 Matthew_Windows
- 15: 928264 Andrew_Symonds
- 16: 3445394 Mark_Ilott
- 17: 6105387 Dominic_Hewson
- 18: 3551407 Tim_Hancock
- 19: 1622322 Gloucestershire_County_Cricket_Club
- 20: 759125 England_cricket_team
- 21: 413176 Graham_Gooch
- 22: 1622252 Essex_County_Cricket_Club
- 23: 5921843 Paul_Prichard
- 24: 1177719 Kent_County_Cricket_Club
- 25: 5882 Cardiff
- 26: 1622320 Glamorgan_County_Cricket_Club
- 27: 766575 Surrey_County_Cricket_Club
- 28: 667341 Nottinghamshire_County_Cricket_Club
- 29: 766575 Surrey_County_Cricket_Club
- 30: 667341 Nottinghamshire_County_Cricket_Club
- 31: 849500 Trent_Bridge
- 32: 26057429 Ali_Brown
- 33: 26057429 Ali_Brown
- 34: 766575 Surrey_County_Cricket_Club
- 35: 1622318 Leicestershire_County_Cricket_Club
- 36: 1919370 Hampshire_County_Cricket_Club
- 37: 51155 Leicester
- 38: 10300919 David_Millns
- 39: 24080381 Gordon_Parsons_(cricketer)
- 40: 3605301 Vince_Wells
- 41: 1919370 Hampshire_County_Cricket_Club
- 42: 869945 Yorkshire_County_Cricket_Club
- 43: 749919 Lancashire_County_Cricket_Club

```
44: 660597 Old_Trafford_Cricket_Ground
45: 869945 Yorkshire_County_Cricket_Club
46: 749919 Lancashire_County_Cricket_Club
47: 413178 Darren_Gough
48: 749919 Lancashire_County_Cricket_Club
49: 1224723 Neil_Fairbrother
1: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.640119,
final_local_scores=0.592649, pem=-2.3459642,
similarity=34.53326]
2: 17867 London
    [pem_pos=0, final_scores=0.57183224,
final_local_scores=0.6517646, pem=-0.09086737,
similarity=24.033817]
3: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.94776016,
final_local_scores=0.92870027, pem=-2.3459642,
similarity=44.5617]
4: 9316 England
    [pem_pos=1, final_scores=0.5373989,
final_local_scores=0.5857056, pem=-1.3527626,
similarity=28.924868]
5: 2856979 Neil_Williams_(cricketer)
    [pem_pos=1, final_scores=0.34977096,
final_local_scores=0.46261537, pem=-0.42566785,
similarity=20.209927]
6: 3445394 Mark_Ilott
    [pem_pos=0, final_scores=0.7303861,
final_local_scores=0.8111974, pem=0.0, similarity=28.29746]
7: 1622322 Gloucestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.651298,
final_local_scores=0.60471886, pem=-1.379793,
similarity=29.63926]
8: 3445394 Mark_Ilott
    [pem_pos=0, final_scores=0.80295163,
final_local_scores=0.8962905, pem=0.0, similarity=30.836807]
9: 1622322 Gloucestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.70076686,
final_local_scores=0.6587561, pem=-1.379793,
similarity=31.251839]
10: 264105 Colchester
    [pem_pos=0, final_scores=0.3617894,
final_local_scores=0.40053147, pem=-0.1350495,
similarity=16.77679]
11: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.87147003,
final_local_scores=0.8453648, pem=-2.3459642,
similarity=42.074802]
12: 1829984 Derbyshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.87466043,
final_local_scores=0.8510141, pem=-1.4828352,
similarity=37.549557]
```

```
13: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.8830361,
final_local_scores=0.85799897, pem=-2.3459642,
similarity=42.45183]
14: 6136803 Matthew_Windows
    [pem_pos=0, final_scores=0.69757986,
final_local_scores=0.77749735, pem=0.0, similarity=27.291784]
15: 928264 Andrew_Symonds
    [pem_pos=0, final_scores=0.9330817,
final_local_scores=1.0011177, pem=0.0, similarity=33.96506]
16: 3445394 Mark_Ilott
    [pem_pos=0, final_scores=0.88775676,
final_local_scores=0.9889272, pem=0.0, similarity=33.601273]
17: 6105387 Dominic_Hewson
    [pem_pos=0, final_scores=0.5174661,
final_local_scores=0.5903806, pem=0.0, similarity=21.707848]
18: 3551407 Tim_Hancock
    [pem_pos=0, final_scores=0.4902305,
final_local_scores=0.5710138, pem=0.0, similarity=21.129904]
19: 1622322 Gloucestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6833901,
final_local_scores=0.6397746, pem=-1.379793,
similarity=30.685394]
20: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.8221255,
final_local_scores=0.8081513, pem=-3.3054674,
similarity=46.182205]
21: 413176 Graham_Gooch
    [pem_pos=0, final_scores=0.93976825,
final_local_scores=0.9788886, pem=0.0, similarity=33.3017]
22: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=1.0256909,
final_local_scores=1.0138278, pem=-2.3459642,
similarity=47.102074]
23: 5921843 Paul_Prichard
    [pem_pos=0, final_scores=0.6632105,
final_local_scores=0.696291, pem=0.0, similarity=24.868425]
24: 1177719 Kent_County_Cricket_Club
    [pem_pos=1, final_scores=0.8046978,
final_local_scores=0.77534086, pem=-2.5096242,
similarity=40.87516]
25: 5882 Cardiff
    [pem_pos=0, final_scores=0.59797525,
final_local_scores=0.6540833, pem=-0.20401932,
similarity=24.718351]
26: 1622320 Glamorgan_County_Cricket_Club
    [pem_pos=1, final_scores=0.8169134,
final_local_scores=0.7905808, pem=-0.7903395,
similarity=31.980211]
27: 766575 Surrey_County_Cricket_Club
    [pem_pos=1, final_scores=0.76972634,
final_local_scores=0.73349684, pem=-1.7952642,
```

```
similarity=35.741646]
28: 667341 Nottinghamshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.5783116,
final_local_scores=0.5252185, pem=-1.1067253,
similarity=25.781824]
29: 766575 Surrey_County_Cricket_Club
    [pem_pos=1, final_scores=0.75286084,
final_local_scores=0.7150739, pem=-1.7952642,
similarity=35.191868]
30: 667341 Nottinghamshire_County_Cricket_Club
    [pem_pos=0, final_scores=0.68182,
final_local_scores=0.6382857, pem=-0.103827626,
similarity=23.70206]
31: 849500 Trent_Bridge
    [pem_pos=0, final_scores=0.6599032,
final_local_scores=0.64571136, pem=0.0, similarity=23.359028]
32: 26057429 Ali_Brown
    [pem_pos=2, final_scores=0.62145805,
final_local_scores=0.6494685, pem=-0.64077926,
similarity=26.955805]
33: 26057429 Ali_Brown
    [pem_pos=2, final_scores=0.57010037,
final_local_scores=0.593368, pem=-0.64077926,
similarity=25.281654]
34: 766575 Surrey_County_Cricket_Club
    [pem_pos=1, final_scores=0.69203895,
final_local_scores=0.6486352, pem=-1.7952642,
similarity=33.209206]
35: 1622318 Leicestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.7435496,
final_local_scores=0.7095954, pem=-1.0465667,
similarity=30.956848]
36: 1919370 Hampshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.7701135,
final_local_scores=0.74026793, pem=-1.9232318,
similarity=36.639618]
37: 51155 Leicester
    [pem_pos=0, final_scores=0.42819148,
final_local_scores=0.4403718, pem=-0.2625849,
similarity=18.659264]
38: 10300919 David_Millns
    [pem_pos=0, final_scores=0.5548431,
final_local_scores=0.63297427, pem=0.0, similarity=22.978928]
39: 24080381 Gordon_Parsons_(cricketer)
    [pem_pos=1, final_scores=0.5902728,
final_local_scores=0.6935357, pem=-1.7917595,
similarity=34.530067]
40: 3605301 Vince_Wells
    [pem_pos=0, final_scores=0.51893187,
final_local_scores=0.601425, pem=0.0, similarity=22.037434]
41: 1919370 Hampshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.85151607,
```

```

final_local_scores=0.8291879, pem=-1.9232318,
similarity=39.293167]
42: 869945 Yorkshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6965954,
final_local_scores=0.65190023, pem=-1.7779728,
similarity=33.21261]
43: 749919 Lancashire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6308052,
final_local_scores=0.5845237, pem=-1.7295884,
similarity=30.938837]
44: 660597 Old_Trafford_Cricket_Ground
    [pem_pos=1, final_scores=0.62441736,
final_local_scores=0.60659343, pem=-0.6437659,
similarity=25.69257]
45: 869945 Yorkshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.7019927,
final_local_scores=0.65779597, pem=-1.7779728,
similarity=33.38855]
46: 749919 Lancashire_County_Cricket_Club
    [pem_pos=1, final_scores=0.67192787,
final_local_scores=0.629444, pem=-1.7295884,
similarity=32.279346]
47: 413178 Darren_Gough
    [pem_pos=0, final_scores=0.86633784,
final_local_scores=0.9346922, pem=0.0, similarity=31.982794]
48: 749919 Lancashire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6564227,
final_local_scores=0.61250687, pem=-1.7295884,
similarity=31.77391]
49: 1224723 Neil_Fairbrother
    [pem_pos=0, final_scores=0.80319995,
final_local_scores=0.87008137, pem=0.0, similarity=30.054678]

```

Sample: b'250_CRICKET'

CRICKET - [1 1 ENGLAND 1] V [2 2 PAKISTAN 2] FINAL TEST
SCOREBOARD . [3 3 LONDON 3] 1996-08-24 Scoreboard on the
third day of the third and final test between [4 4 England 4]
and [5 5 Pakistan 5] at [6 6 The Oval 6] on Saturday : [7 7 En
gland 7] first innings 326 (J. Crawley 106 , G. Thorpe 54 ; [
8 8 Waqar Younis 8] 4-95) [9 9 Pakistan 9] first innings
(overnight 229-1) [10 10 Saeed Anwar 10] c [11 11 Croft 11]
b [12 12 Cork 12] 176 Aamir Sohail c [13 13 Cork 13] b [14 14
Croft 14] 46 [15 15 Ijaz Ahmed 15] c 1 Stewart b [16 16 Mulla
lly 16] 61 [17 17 Inzamam-ul-Haq 17] c [18 18 Hussain 18] b [1
9 19 Mullally 19] 35 [20 20 Salim Malik 20] not out 2 [21 21 A
sif Mujtaba 21] not out 1 Extras (b-4 lb-3 nb-11) 18 Total
(for four wickets) 339 Fall of wickets : 1-106 2-239 3-334
4-334 To bat : [22 22 Wasim Akram 22] , [23 23 Moin Khan 23] ,
[24 24 Mushtaq Ahmed 24] , [25 25 Waqar Younis 25] ,
Mohammad Akam Bowling (to date) : Lewis 12-1-76-0 , [26 26 M
ullally 26] 22-6-56-2 , [27 27 Croft 27] 29-6-64-1 , [28 28 Co
rk 28] 14.3-4-45-1 , 2 Salisbury 17-0-91-0

```
1: 759125 England_cricket_team
2: 759295 Pakistan_national_cricket_team
3: 17867 London
4: 759125 England_cricket_team
5: 759295 Pakistan_national_cricket_team
6: 387687 The_Oval
7: 759125 England_cricket_team
8: 894433 Waqar_Younis
9: 759295 Pakistan_national_cricket_team
10: 894437 Saeed_Anwar
11: 1160680 Robert_Croft
12: 2036840 Dominic_Cork
13: 2036840 Dominic_Cork
14: 1160680 Robert_Croft
15: 938863 Ijaz_Ahmed_(cricketer)
16: 2222543 Alan_Mullally
17: 896403 Inzamam-ul-Haq
18: 682925 Nasser_Hussain
19: 2222543 Alan_Mullally
20: 938861 Saleem_Malik
21: 3976606 Asif_Mujtaba
22: 501393 Wasim_Akram
23: 554132 Moin_Khan
24: 896405 Mushtaq_Ahmed
25: 894433 Waqar_Younis
26: 2222543 Alan_Mullally
27: 1160680 Robert_Croft
28: 2036840 Dominic_Cork
1: 896414 Alec_Stewart
    [winner: 876610 Dave_Stewart_(baseball), pem_pos=14,
final_scores=-0.73052627, final_local_scores=-0.67802477,
pem=-4.102782, similarity=6.1676664|
     RECALL MISS: 896414 Alec_Stewart]
2: 2416655 Ian_Salisbury
    [winner: 72119 Salisbury, pem_pos=0,
final_scores=-0.45350865, final_local_scores=-0.43754697,
pem=-0.80015767, similarity=-4.6161795|
     RECALL MISS: 2416655 Ian_Salisbury]
1: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.6398214,
final_local_scores=0.60503536, pem=-3.3054674,
similarity=40.12082]
2: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7176525,
final_local_scores=0.693801, pem=-2.7556365,
similarity=39.779697]
3: 17867 London
    [pem_pos=0, final_scores=0.554136,
final_local_scores=0.6760777, pem=-0.09086737,
similarity=24.75937]
4: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.81709105,
```

```
final_local_scores=0.7986755, pem=-3.3054674,
similarity=45.89943]
5: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.85047096,
final_local_scores=0.838885, pem=-2.7556365,
similarity=44.109295]
6: 387687 The_Oval
    [pem_pos=0, final_scores=0.62158936,
final_local_scores=0.61662155, pem=0.0, similarity=22.49093]
7: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.81846327,
final_local_scores=0.8001744, pem=-3.3054674,
similarity=45.94416]
8: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.6969045,
final_local_scores=0.6944665, pem=0.0, similarity=24.813978]
9: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.8547608,
final_local_scores=0.843571, pem=-2.7556365,
similarity=44.249134]
10: 894437 Saeed_Anwar
    [pem_pos=0, final_scores=0.79190904,
final_local_scores=0.8060016, pem=0.0, similarity=28.142408]
11: 1160680 Robert_Croft
    [pem_pos=7, final_scores=1.0033631,
final_local_scores=1.0444374, pem=-3.16984,
similarity=52.49589]
12: 2036840 Dominic_Cork
    [pem_pos=8, final_scores=0.91555005,
final_local_scores=0.9257632, pem=-3.606134,
similarity=51.32705]
13: 2036840 Dominic_Cork
    [pem_pos=8, final_scores=1.143958,
final_local_scores=1.1752641, pem=-3.606134,
similarity=58.77266]
14: 1160680 Robert_Croft
    [pem_pos=7, final_scores=0.99930805,
final_local_scores=1.0400078, pem=-3.16984,
similarity=52.36371]
15: 938863 Ijaz_Ahmed_(cricketer)
    [pem_pos=0, final_scores=0.6174442,
final_local_scores=0.6198503, pem=0.0, similarity=22.587282]
16: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.7675989,
final_local_scores=0.7883271, pem=-0.73930055,
similarity=31.635399]
17: 896403 Inzamam-ul-Haq
    [pem_pos=0, final_scores=1.024642,
final_local_scores=1.051431, pem=0.0, similarity=35.46651]
18: 682925 Nasser_Hussain
    [pem_pos=3, final_scores=0.6867096,
final_local_scores=0.68914217, pem=-2.447519,
```

```

similarity=37.965076]
19: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.83207554,
final_local_scores=0.8587581, pem=-0.73930055,
similarity=33.737198]
20: 938861 Saleem_Malik
    [pem_pos=0, final_scores=0.7426274,
final_local_scores=0.77403015, pem=0.0, similarity=27.188316]
21: 3976606 Asif_Mujtaba
    [pem_pos=0, final_scores=0.76488966,
final_local_scores=0.80528295, pem=0.0, similarity=28.120962]
22: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.75463897,
final_local_scores=0.75295836, pem=0.0, similarity=26.559492]
23: 554132 Moin_Khan
    [pem_pos=0, final_scores=0.7412184,
final_local_scores=0.7891416, pem=0.0, similarity=27.639273]
24: 896405 Mushtaq_Ahmed
    [pem_pos=0, final_scores=0.63155264,
final_local_scores=0.63553226, pem=0.0, similarity=23.055264]
25: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.73875046,
final_local_scores=0.74017686, pem=0.0, similarity=26.178066]
26: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.69995886,
final_local_scores=0.7144407, pem=-0.73930055,
similarity=29.430481]
27: 1160680 Robert_Croft
    [pem_pos=7, final_scores=0.85664743,
final_local_scores=0.88417274, pem=-3.16984,
similarity=47.713272]
28: 2036840 Dominic_Cork
    [pem_pos=8, final_scores=0.8737269,
final_local_scores=0.8800778, pem=-3.606134,
similarity=49.963707]

```

Sample: b'251_CRICKET'

CRICKET - [1 1 ENGLISH COUNTY CHAMPIONSHIP 1] SCORES . [2 2 L
ONDON 2] 1996-08-24 Close of play scores in four-day [3 3 Eng
lish County Championship 3] cricket matches on Saturday :
Final day At [4 4 Weston-super-Mare 4] : Match abandoned as a
draw - rain . [5 5 Durham 5] 326 (D. Cox 95 not out , S.
Campbell 69 ; G. Rose 7-73) . [6 6 Somerset 6] 298-6 (M.
Lathwell 85 , R. Harden 65) . [7 7 Somerset 7] 9 points , [8
8 Durham 8] 8 . Third day At [9 9 Colchester 9] : [10 10 Glou
cestershire 10] 280 and 27-4 . [11 11 Essex 11] 532-8
declared (G. Gooch 111 , R. Irani 91 , P. Prichard 88 , D.
Robinson 72 ; M. Alleyne 4-80) . At [12 12 Cardiff 12] : [13
13 Kent 13] 255-3 (D. Fulton 64 , M. Walker 59 , C. Hooper
52 not out) v [14 14 Glamorgan 14] . No play - rain . At [15
15 Leicester 15] : [16 16 Leicestershire 16] 353 (P. Simmons
108 , P. Nixon 67 ; S. Renshaw 4-56 , J. Bovill 4-102) . [17]

17 Hampshire 17] 81-7 . At [18 18 Northampton 18] : [19 19 Sussex 19] 389 and 112 (C. Ambrose 6-26) . [20 20 Northamptonshire 20] 361 (K. Curran 117 , D. Ripley 66 not out) and 42-3 . At [21 21 Trent Bridge 21] : [22 22 Nottinghamshire 22] 446-9 declared (G. Archer 143 , M. Dowman 107 , W. Noon 57 ; B. Julian 4-104) . [23 23 Surrey 23] 128-4 (A. Brown 56 not out) . At [24 24 Worcester 24] : [25 25 Warwickshire 25] 310 (A. Giles 83 , T. Munton 54 not out , W. Khan 52 ; R. Illingworth 4-54 , S. Lampitt 4-90) . [26 26 Worcestershire 26] 205-9 (K. Spiring 52) . At [27 27 Headingley 27] : [28 28 Yorkshire 28] 529-8 declared (C. White 181 , R. Blakey 109 not out , M. Moxon 66 , M. Vaughan 57) . [29 29 Lancashire 29] 323 (N. Fairbrother 86 , [30 30 M. Watkinson 30] 64 ; [31 31 D. Gough 31] 4-53) and 210-5 (N. Speak 65 not out , N. Fairbrother 55) .

1: 683599 County_Championship
2: 17867 London
3: 683599 County_Championship
4: 330072 Weston-super-Mare
5: 1176371 Durham_County_Cricket_Club
6: 1622178 Somerset_County_Cricket_Club
7: 1622178 Somerset_County_Cricket_Club
8: 1176371 Durham_County_Cricket_Club
9: 11568817 Castle_Park_Cricket_Ground
10: 1622322 Gloucestershire_County_Cricket_Club
11: 1622252 Essex_County_Cricket_Club
12: 5882 Cardiff
13: 1177719 Kent_County_Cricket_Club
14: 1622320 Glamorgan_County_Cricket_Club
15: 51155 Leicester
16: 1622318 Leicestershire_County_Cricket_Club
17: 1919370 Hampshire_County_Cricket_Club
18: 63734 Northampton
19: 1763691 Sussex_County_Cricket_Club
20: 530308 Northamptonshire_County_Cricket_Club
21: 849500 Trent_Bridge
22: 667341 Nottinghamshire_County_Cricket_Club
23: 766575 Surrey_County_Cricket_Club
24: 58681 Worcester
25: 667292 Warwickshire_County_Cricket_Club
26: 1156428 Worcestershire_County_Cricket_Club
27: 1876060 Headingley_Stadium
28: 869945 Yorkshire_County_Cricket_Club
29: 749919 Lancashire_County_Cricket_Club
30: 2625808 Mike_Watkinson
31: 413178 Darren_Gough
1: 683599 County_Championship
 [pem_pos=0, final_scores=0.37543494,
final_local_scores=0.3199294, pem=0.0, similarity=13.637044]
2: 17867 London
 [pem_pos=0, final_scores=0.6787261,
final_local_scores=0.7464871, pem=-0.09086737,

```
similarity=26.860527]
3: 683599 County_Championship
    [pem_pos=0, final_scores=0.23962854,
final_local_scores=0.1715816, pem=0.0, similarity=9.21005]
4: 330072 Weston-super-Mare
    [pem_pos=0, final_scores=0.32384127,
final_local_scores=0.34634736, pem=0.0, similarity=14.425409]
5: 1176371 Durham_County_Cricket_Club
    [pem_pos=2, final_scores=0.86666733,
final_local_scores=0.85569435, pem=-2.3952909,
similarity=42.651306]
6: 1622178 Somerset_County_Cricket_Club
    [pem_pos=1, final_scores=0.6572686,
final_local_scores=0.6113142, pem=-2.4139037,
similarity=35.459732]
7: 1622178 Somerset_County_Cricket_Club
    [pem_pos=1, final_scores=0.73342603,
final_local_scores=0.6945046, pem=-2.4139037,
similarity=37.9423]
8: 1176371 Durham_County_Cricket_Club
    [pem_pos=2, final_scores=0.9350706,
final_local_scores=0.9304145, pem=-2.3952909,
similarity=44.881104]
9: 11568817 Castle_Park_Cricket_Ground
    [pem_pos=17, final_scores=0.3036595,
final_local_scores=0.36781138, pem=-2.976298,
similarity=31.251509]
10: 1622322 Gloucestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.7114143,
final_local_scores=0.66998154, pem=-1.379793,
similarity=31.586828]
11: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.8556029,
final_local_scores=0.83178675, pem=-2.3459642,
similarity=41.669605]
12: 5882 Cardiff
    [pem_pos=0, final_scores=0.6290024,
final_local_scores=0.66141474, pem=-0.20401932,
similarity=24.937136]
13: 1177719 Kent_County_Cricket_Club
    [pem_pos=1, final_scores=0.764304,
final_local_scores=0.7295304, pem=-2.5096242,
similarity=39.508083]
14: 1622320 Glamorgan_County_Cricket_Club
    [pem_pos=1, final_scores=0.6584827,
final_local_scores=0.6147187, pem=-0.7903395,
similarity=26.732134]
15: 51155 Leicester
    [pem_pos=0, final_scores=0.6294672,
final_local_scores=0.63709855, pem=-0.2625849,
similarity=24.529982]
16: 1622318 Leicestershire_County_Cricket_Club
```

```
[pem_pos=1, final_scores=0.7738001,
final_local_scores=0.7445918, pem=-1.0465667,
similarity=32.00121]
17: 1919370 Hampshire_County_Cricket_Club
[pem_pos=1, final_scores=0.68853897,
final_local_scores=0.65160924, pem=-1.9232318,
similarity=33.993866]
18: 63734 Northampton
[pem_pos=0, final_scores=0.53556967,
final_local_scores=0.5406135, pem=-0.5110059,
similarity=23.001627]
19: 1763691 Sussex_County_Cricket_Club
[pem_pos=1, final_scores=0.6632466,
final_local_scores=0.6328704, pem=-1.5566913,
similarity=31.441357]
20: 530308 Northamptonshire_County_Cricket_Club
[pem_pos=1, final_scores=0.40056306,
final_local_scores=0.3372295, pem=-1.2069724,
similarity=20.717016]
21: 849500 Trent_Bridge
[pem_pos=0, final_scores=0.6319717,
final_local_scores=0.61692995, pem=0.0, similarity=22.500134]
22: 667341 Nottinghamshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6763441,
final_local_scores=0.6318853, pem=-1.1067253,
similarity=28.964975]
23: 766575 Surrey_County_Cricket_Club
[pem_pos=1, final_scores=0.6240719,
final_local_scores=0.5726053, pem=-1.7952642,
similarity=30.940325]
24: 58681 Worcester
[pem_pos=0, final_scores=0.52589643,
final_local_scores=0.5488828, pem=-0.59997296,
similarity=23.732216]
25: 667292 Warwickshire_County_Cricket_Club
[pem_pos=1, final_scores=0.7669757,
final_local_scores=0.73117137, pem=-1.1093602,
similarity=31.942198]
26: 1156428 Worcestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.6304324,
final_local_scores=0.58451587, pem=-0.9164197,
similarity=26.516464]
27: 1876060 Headingley_Stadium
[pem_pos=0, final_scores=0.23803668,
final_local_scores=0.21118125, pem=-0.30334172,
similarity=12.041402]
28: 869945 Yorkshire_County_Cricket_Club
[pem_pos=1, final_scores=0.71349615,
final_local_scores=0.6708762, pem=-1.7779728,
similarity=33.77889]
29: 749919 Lancashire_County_Cricket_Club
[pem_pos=1, final_scores=0.5730756,
```

```
final_local_scores=0.5197752, pem=-1.7295884,  
similarity=29.006613]  
30: 2625808 Mike_Watkinson  
    [pem_pos=0, final_scores=0.42058128,  
final_local_scores=0.47143728, pem=0.0, similarity=18.158342]  
31: 413178 Darren_Gough  
    [pem_pos=0, final_scores=0.48662484,  
final_local_scores=0.5376586, pem=0.0, similarity=20.134518]
```

Sample: b'252_SOCCER'

SOCCER - [1 1 SCOTTISH 1] LEAGUE RESULTS . [2 2 GLASGOW 2]
1996-08-24 Results of [3 3 Scottish 3] league soccer matches
on Saturday : Premier division [4 4 Hibernian 4] 0 [5 5 Dunfermline 5] 0 [6 6 Kilmarnock 6] 1 [7 7 Celtic 7] 3 [8 8 Raith 8] 0 [9 9 Motherwell 9] 3 [10 10 Rangers 10] 1 [11 11 Dundee United 11] 0 Playing Sunday : [12 12 Aberdeen 12] v [13 13 Hearts 13] Division one [14 14 Airdrieonians 14] 0 [15 15 East Fife 15] 0 [16 16 Clydebank 16] 1 1 Stirling 0 [17 17 Dundee 17] 2 [18 18 Greenock Morton 18] 1 [19 19 Falkirk 19] 1 [20 20 Partick 20] 0 [21 21 St Mirren 21] 0 [22 22 St Johnstone 22] 3
Division two 2 Berwick 0 [23 23 Stenhousemuir 23] 6 [24 24 Brechin 24] 1 [25 25 Ayr 25] 1 3 Hamilton 2 [26 26 Clyde 26] 0 [27 27 Queen of the South 27] 2 [28 28 Dumbarton 28] 1 [29 29 Stranraer 29] 1 [30 30 Livingston 30] 2 Division three [31 31 Alloa 31] 1 [32 32 Arbroath 32] 1 [33 33 Cowdenbeath 33] 1
Monstrose 0 [34 34 Forfar 34] 3 [35 35 Inverness 35] 1 [36 36 Ross County 36] 1 Queen 's Park 2 Played Friday : [37 37 East Stirling 37] 0 Albion 1

1: 26994 Scotland
2: 68736 Glasgow
3: 26994 Scotland
4: 176747 Hibernian_F.C.
5: 451181 Dunfermline_Athletic_F.C.
6: 405663 Kilmarnock_F.C.
7: 69708 Celtic_F.C.
8: 880348 Raith_Rovers_F.C.
9: 215398 Motherwell_F.C.
10: 69713 Rangers_F.C.
11: 331074 Dundee_United_F.C.
12: 237019 Aberdeen_F.C.
13: 302212 Heart_of_Midlothian_F.C.
14: 1216952 Airdrieonians_F.C._(1878)
15: 451245 East_Fife_F.C.
16: 1347678 Clydebank_F.C.
17: 306504 Dundee_F.C.
18: 367890 Greenock_Morton_F.C.
19: 560087 Falkirk_F.C.
20: 226415 Partick_Thistle_F.C.
21: 783089 St._Mirren_F.C.
22: 393756 St._Johnstone_F.C.
23: 653922 Stenhousemuir_F.C.
24: 451241 Brechin_City_F.C.

25: 451240 Ayr_United_F.C.
26: 451243 Clyde_F.C.
27: 880354 Queen_of_the_South_F.C.
28: 443155 Dumbarton_F.C.
29: 880391 Stranraer_F.C.
30: 451183 Livingston_F.C.
31: 451237 Alloa_Athletic_F.C.
32: 451238 Arbroath_F.C.
33: 451244 Cowdenbeath_F.C.
34: 408985 Forfar_Athletic_F.C.
35: 743540 Inverness_Caledonian_Thistle_F.C.
36: 880380 Ross_County_F.C.
37: 817230 East_Stirlingshire_F.C.
1: 880388 Stirling_Albion_F.C.
[winner: 61311 Stirling, pem_pos=0,
final_scores=-0.73179704, final_local_scores=-0.7977853,
pem=-0.2433456, similarity=-18.39444]
RECALL MISS: 880388 Stirling_Albion_F.C.]
2: 451242 Berwick_Rangers_F.C.
[winner: 132088 Berwick,_Pennsylvania, pem_pos=0,
final_scores=-0.79336435, final_local_scores=-0.76844996,
pem=-1.6255274, similarity=-10.002496]
RECALL MISS: 451242 Berwick_Rangers_F.C.]
3: 309502 Hamilton_Academical_F.C.
[winner: 98626 Hamilton,_South_Lanarkshire, pem_pos=2,
final_scores=-0.94375837, final_local_scores=-1.0366769,
pem=-3.0741625, similarity=-10.129019]
RECALL MISS: 309502 Hamilton_Academical_F.C.]
1: 26994 Scotland
[pem_pos=0, final_scores=0.5156608,
final_local_scores=0.56250155, pem=-0.47951394,
similarity=23.483553]
2: 68736 Glasgow
[pem_pos=0, final_scores=0.48785383,
final_local_scores=0.5173497, pem=-0.15604196,
similarity=20.37704]
3: 26994 Scotland
[pem_pos=0, final_scores=0.5088472,
final_local_scores=0.5550587, pem=-0.47951394,
similarity=23.261444]
4: 176747 Hibernian_F.C.
[pem_pos=0, final_scores=0.7669253,
final_local_scores=0.7250304, pem=0.0, similarity=25.726067]
5: 451181 Dunfermline_Athletic_F.C.
[pem_pos=1, final_scores=0.6042843,
final_local_scores=0.5478057, pem=-1.9918579,
similarity=31.269361]
6: 405663 Kilmarnock_F.C.
[pem_pos=1, final_scores=0.7138191,
final_local_scores=0.6707782, pem=-1.4326576,
similarity=31.898085]
7: 69708 Celtic_F.C.

[pem_pos=0, final_scores=0.6922992,
final_local_scores=0.6599453, pem=-0.9785444,
similarity=29.105272]
8: 880348 Raith_Rovers_F.C.
[pem_pos=0, final_scores=0.6302411,
final_local_scores=0.5803988, pem=0.0, similarity=21.409971]
9: 215398 Motherwell_F.C.
[pem_pos=0, final_scores=0.66760117,
final_local_scores=0.6184871, pem=-0.3361376,
similarity=24.37457]
10: 69713 Rangers_F.C.
[pem_pos=0, final_scores=0.59852266,
final_local_scores=0.5541243, pem=-0.30191645,
similarity=22.267756]
11: 331074 Dundee_United_F.C.
[pem_pos=0, final_scores=0.42889017,
final_local_scores=0.35990688, pem=0.0, similarity=14.830052]
12: 237019 Aberdeen_F.C.
[pem_pos=1, final_scores=0.93924946,
final_local_scores=0.91772515, pem=-1.4493513,
similarity=39.35826]
13: 302212 Heart_of_Midlothian_F.C.
[pem_pos=0, final_scores=0.73614347,
final_local_scores=0.6960744, pem=-0.29860955,
similarity=26.485847]
14: 1216952 Airdrieonians_F.C._(1878)
[pem_pos=1, final_scores=0.23505615,
final_local_scores=0.1570212, pem=-1.8786064,
similarity=18.991692]
15: 451245 East_Fife_F.C.
[pem_pos=0, final_scores=0.96019024,
final_local_scores=0.9492236, pem=0.0, similarity=32.41644]
16: 1347678 Clydebank_F.C.
[pem_pos=1, final_scores=0.52674204,
final_local_scores=0.48229623, pem=-1.3275584,
similarity=25.701864]
17: 306504 Dundee_F.C.
[pem_pos=1, final_scores=0.68725544,
final_local_scores=0.639314, pem=-1.3661808,
similarity=30.597622]
18: 367890 Greenock_Morton_F.C.
[pem_pos=0, final_scores=0.69955415,
final_local_scores=0.65773857, pem=0.0, similarity=23.717943]
19: 560087 Falkirk_F.C.
[pem_pos=1, final_scores=0.62384945,
final_local_scores=0.5717181, pem=-1.2052568,
similarity=27.705296]
20: 226415 Partick_Thistle_F.C.
[pem_pos=2, final_scores=0.7143107,
final_local_scores=0.6700516, pem=-1.2851555,
similarity=31.074265]
21: 783089 St._Mirren_F.C.

```
[pem_pos=0, final_scores=0.67171293,
final_local_scores=0.6295791, pem=0.0, similarity=22.87761]
22: 393756 St._Johnstone_F.C.
    [pem_pos=0, final_scores=0.78473836,
final_local_scores=0.7446634, pem=0.0, similarity=26.311954]
23: 653922 Stenhousemuir_F.C.
    [pem_pos=0, final_scores=0.6899827,
final_local_scores=0.6595212, pem=0.0, similarity=23.771141]
24: 451241 Brechin_City_F.C.
    [pem_pos=2, final_scores=0.71555305,
final_local_scores=0.6829254, pem=-1.9178002,
similarity=34.898865]
25: 451240 Ayr_United_F.C.
    [pem_pos=17, final_scores=0.57603186,
final_local_scores=0.5214435, pem=-3.340128,
similarity=37.814762]
26: 451243 Clyde_F.C.
    [pem_pos=0, final_scores=0.97479934,
final_local_scores=0.9546643, pem=-0.9832426,
similarity=37.925823]
27: 880354 Queen_of_the_South_F.C.
    [pem_pos=0, final_scores=0.62496907,
final_local_scores=0.57837903, pem=0.0, similarity=21.349697]
28: 443155 Dumbarton_F.C.
    [pem_pos=1, final_scores=0.55950177,
final_local_scores=0.51353294, pem=-1.254432,
similarity=26.236357]
29: 880391 Stranraer_F.C.
    [pem_pos=1, final_scores=0.6636421,
final_local_scores=0.6272946, pem=-1.0295545,
similarity=28.408312]
30: 451183 Livingston_F.C.
    [pem_pos=0, final_scores=1.1004428,
final_local_scores=1.1021755, pem=-1.8614511,
similarity=47.103687]
31: 451237 Alloa_Athletic_F.C.
    [pem_pos=1, final_scores=0.74566114,
final_local_scores=0.7147749, pem=-0.79396856,
similarity=29.737747]
32: 451238 Arbroath_F.C.
    [pem_pos=1, final_scores=0.9239941,
final_local_scores=0.90732104, pem=-0.8612016,
similarity=35.84933]
33: 451244 Cowdenbeath_F.C.
    [pem_pos=0, final_scores=0.6480462,
final_local_scores=0.6110901, pem=-0.24741817,
similarity=23.671358]
34: 408985 Forfar_Athletic_F.C.
    [pem_pos=1, final_scores=0.68158644,
final_local_scores=0.64814955, pem=-1.7646968,
similarity=33.028484]
35: 743540 Inverness_Caledonian_Thistle_F.C.
```

```
[pem_pos=6, final_scores=0.9084291,
final_local_scores=0.89381796, pem=-3.2290676,
similarity=48.323196]
36: 880380 Ross_County_F.C.
    [pem_pos=0, final_scores=0.5065583,
final_local_scores=0.45233226, pem=0.0, similarity=17.58821]
37: 817230 East_Stirlingshire_F.C.
    [pem_pos=0, final_scores=0.6856946,
final_local_scores=0.65794486, pem=0.0, similarity=23.7241]
```

Sample: b'253_SOCCER'

SOCCER - OUT-OF-SORTS [1 1 NEWCASTLE 1] CRASH 2-1 AT HOME . [2 2 LONDON 2] 1996-08-24 [3 3 Newcastle 3] 's early season teething problems continued on Saturday when they lost 2-1 at home to premier league pacesetters [4 4 Sheffield Wednesday 4] . [5 5 England 5] striker [6 6 Alan Shearer 6] gave [7 7 Kevin Keegan 7] 's talent-laden side the lead from the penalty spot after 13 minutes after Wednesday 's [8 8 Yugoslav 8] [9 9 Dejan Stefanovic 9] pulled down [10 10 Colombian 10] forward [11 11 Faustino Asprilla 11] . But two minutes later [12 12 Wednesday 12] equalised through [13 13 Peter Atherton 13] , who found space in the penalty area to meet [14 14 Mark Pembridge 14] 's free kick with a precise glancing header . [15 15 Guy Whittingham 15] stole three points for the [16 16 Yorkshire 16] side with a goal 10 minutes from time . To add to [17 17 Newcastle 17] 's misery , [18 18 England 18] striker [19 19 Les Ferdinand 19] was stretchered off in the second half . [20 20 Wednesday 20] , who escaped relegation on the final day of last season , have now won their first three games of the season . Elsewhere , title hopefuls [21 21 Liverpool 21] were held 0-0 at home by newly-promoted [22 22 Sunderland 22] , and in [23 23 London 23] , the tie between [24 24 Tottenham Hotspur 24] and [25 25 Everton 25] also ended goalless . [26 26 Frenchman 26] [27 27 Frank LeBoeuf 27] and [28 28 Italian 28] [29 29 Gianluca Vialli 29] scored their first premier league goals as [30 30 Chelsea 30] beat [31 31 Coventry 31] 2-0 , and managerless [32 32 Arsenal 32] won by the same scoreline at [33 33 Leicester 33] . Last season 's league and [34 34 Cup 34] winners [35 35 Manchester United 35] host 1995 champions [36 36 Blackburn 36] on Sunday .

1: 57802 Newcastle_United_F.C.
2: 17867 London
3: 57802 Newcastle_United_F.C.
4: 177538 Sheffield_Wednesday_F.C.
5: 9904 England_national_football_team
6: 58682 Alan_Shearer
7: 346979 Kevin_Keegan
8: 34244 Yugoslavia
9: 1004542 Dejan_Stefanović
10: 5222 Colombia
11: 1041521 Faustino_Asprilla
12: 177538 Sheffield_Wednesday_F.C.

13: 4249924 Peter_Atherton
14: 1053382 Mark_Pembridge
15: 2592473 Guy_Whittingham
16: 36637 Yorkshire
17: 57802 Newcastle_United_F.C.
18: 9904 England_national_football_team
19: 526001 Les_Ferdinand
20: 177538 Sheffield_Wednesday_F.C.
21: 18119 Liverpool_F.C.
22: 184474 Sunderland_A.F.C.
23: 17867 London
24: 68198 Tottenham_Hotspur_F.C.
25: 91155 Everton_F.C.
26: 5843419 France
27: 1509963 Frank_Leboeuf
28: 14532 Italy
29: 932030 Gianluca_Vialli
30: 7473 Chelsea_F.C.
31: 435021 Coventry_City_F.C.
32: 2174 Arsenal_F.C.
33: 298602 Leicester_City_F.C.
34: 11237 FA_Cup
35: 19961 Manchester_United_F.C.
36: 298599 Blackburn_Rovers_F.C.
1: 57802 Newcastle_United_F.C.
[pem_pos=2, final_scores=0.6624362,
final_local_scores=0.6126658, pem=-2.4885974,
similarity=35.90626]
2: 17867 London
[pem_pos=0, final_scores=0.5752602,
final_local_scores=0.628605, pem=-0.09086737,
similarity=23.34269]
3: 57802 Newcastle_United_F.C.
[pem_pos=2, final_scores=0.67473614,
final_local_scores=0.6261016, pem=-2.4885974,
similarity=36.307213]
4: 177538 Sheffield_Wednesday_F.C.
[pem_pos=0, final_scores=0.3464741,
final_local_scores=0.29227045, pem=0.0, similarity=12.811646]
5: 9904 England_national_football_team
[pem_pos=1, final_scores=0.5132987,
final_local_scores=0.46484983, pem=-2.776464,
similarity=33.0606]
6: 58682 Alan_Shearer
[pem_pos=0, final_scores=0.66825354,
final_local_scores=0.6633575, pem=0.0, similarity=23.885624]
7: 346979 Kevin_Keegan
[pem_pos=0, final_scores=0.4818765,
final_local_scores=0.4651546, pem=0.0, similarity=17.970854]
8: 34244 Yugoslavia
[pem_pos=0, final_scores=0.27028644,
final_local_scores=0.33858013, pem=-0.7596219,

```
similarity=18.324562]
9: 1004542 Dejan_Stefanović
    [pem_pos=0, final_scores=0.4531982,
final_local_scores=0.5183529, pem=0.0, similarity=19.558397]
10: 5222 Colombia
    [pem_pos=0, final_scores=0.46014088,
final_local_scores=0.5744578, pem=-0.32036036,
similarity=22.97485]
11: 1041521 Faustino_Asprilla
    [pem_pos=0, final_scores=0.6689854,
final_local_scores=0.7037393, pem=0.0, similarity=25.090696]
12: 177538 Sheffield_Wednesday_F.C.
    [pem_pos=2, final_scores=0.49353075,
final_local_scores=0.4529075, pem=-1.7182498,
similarity=26.949486]
13: 4249924 Peter_Atherton
    [pem_pos=0, final_scores=0.70517087,
final_local_scores=0.78325796, pem=0.0, similarity=27.463694]
14: 1053382 Mark_Pembridge
    [pem_pos=0, final_scores=0.5508968,
final_local_scores=0.58127636, pem=0.0, similarity=21.43616]
15: 2592473 Guy_Whittingham
    [pem_pos=0, final_scores=0.57827073,
final_local_scores=0.6050985, pem=0.0, similarity=22.147058]
16: 36637 Yorkshire
    [pem_pos=0, final_scores=0.2709399,
final_local_scores=0.300272, pem=-0.31620604,
similarity=14.770005]
17: 57802 Newcastle_United_F.C.
    [pem_pos=2, final_scores=0.7486958,
final_local_scores=0.7068913, pem=-2.4885974,
similarity=38.71814]
18: 9904 England_national_football_team
    [pem_pos=1, final_scores=0.4073829,
final_local_scores=0.34915292, pem=-2.776464,
similarity=29.607975]
19: 526001 Les_Ferdinand
    [pem_pos=0, final_scores=0.54697144,
final_local_scores=0.5441428, pem=0.0, similarity=20.328018]
20: 177538 Sheffield_Wednesday_F.C.
    [pem_pos=2, final_scores=0.5335023,
final_local_scores=0.4965703, pem=-1.7182498,
similarity=28.25247]
21: 18119 Liverpool_F.C.
    [pem_pos=1, final_scores=0.64576614,
final_local_scores=0.59610456, pem=-1.1702383,
similarity=28.2426]
22: 184474 Sunderland_A.F.C.
    [pem_pos=0, final_scores=0.5997631,
final_local_scores=0.55030733, pem=-0.39640337,
similarity=22.667685]
23: 17867 London
```

```
[pem_pos=0, final_scores=0.7216138,
final_local_scores=0.7884741, pem=-0.09086737,
similarity=28.113503]
24: 68198 Tottenham_Hotspur_F.C.
    [pem_pos=0, final_scores=0.5206421,
final_local_scores=0.4600833, pem=0.0, similarity=17.819517]
25: 91155 Everton_F.C.
    [pem_pos=0, final_scores=0.52313733,
final_local_scores=0.46137106, pem=-0.17171584,
similarity=18.791763]
26: 5843419 France
    [pem_pos=1, final_scores=0.4255851,
final_local_scores=0.50085765, pem=-0.7184992,
similarity=22.943615]
27: 1509963 Frank_Leboeuf
    [pem_pos=0, final_scores=0.7007207,
final_local_scores=0.75974727, pem=0.0, similarity=26.762087]
28: 14532 Italy
    [pem_pos=0, final_scores=0.63329893,
final_local_scores=0.751043, pem=-0.75948656,
similarity=30.632542]
29: 932030 Gianluca_Vialli
    [pem_pos=0, final_scores=0.5892151,
final_local_scores=0.6000034, pem=0.0, similarity=21.995012]
30: 7473 Chelsea_F.C.
    [pem_pos=0, final_scores=0.7635372,
final_local_scores=0.7321999, pem=-0.51707125,
similarity=28.751932]
31: 435021 Coventry_City_F.C.
    [pem_pos=2, final_scores=0.6667983,
final_local_scores=0.6306533, pem=-3.1334689,
similarity=39.949955]
32: 2174 Arsenal_F.C.
    [pem_pos=0, final_scores=0.53184086,
final_local_scores=0.4735387, pem=-0.15820743,
similarity=19.08141]
33: 298602 Leicester_City_F.C.
    [pem_pos=3, final_scores=0.7530624,
final_local_scores=0.72203076, pem=-3.2133117,
similarity=43.11104]
34: 11237 FA_Cup
    [pem_pos=6, final_scores=0.23084073,
final_local_scores=0.17556661, pem=-3.1209865,
similarity=26.30138]
35: 19961 Manchester_United_F.C.
    [pem_pos=0, final_scores=0.47470987,
final_local_scores=0.4110936, pem=0.0, similarity=16.357565]
36: 298599 Blackburn_Rovers_F.C.
    [pem_pos=1, final_scores=0.6120551,
final_local_scores=0.56156653, pem=-2.734066,
similarity=35.716255]
```

Sample: b'254_SOCCER'

SOCCER - [1 1 ENGLISH 1] LEAGUE RESULTS . [2 2 LONDON 2]
1996-08-24 Results of [3 3 English 3] league soccer matches
on Saturday : Premier league [4 4 Aston Villa 4] 2 [5 5 Derby
5] 0 [6 6 Chelsea 6] 2 [7 7 Coventry 7] 0 [8 8 Leicester 8] 0
[9 9 Arsenal 9] 2 [10 10 Liverpool 10] 0 [11 11 Sunderland 11]
0 [12 12 Newcastle 12] 1 [13 13 Sheffield Wednesday 13] 2 [14
14 Nottingham Forest 14] 1 [15 15 Middlesbrough 15] 1 [16 16 T
ottenham 16] 0 [17 17 Everton 17] 0 [18 18 West Ham 18] 2 [19
19 Southampton 19] 1 Playing Sunday : [20 20 Manchester United
20] v [21 21 Blackburn 21] Playing Monday : [22 22 Leeds 22]
v [23 23 Wimbledon 23] Division one [24 24 Bolton 24] 3 [25 25
Norwich 25] 1 [26 26 Charlton 26] 1 [27 27 West Bromwich 27]
1 [28 28 Crystal Palace 28] 3 [29 29 Oldham 29] 1 [30 30 Ipsw
ich 30] 5 [31 31 Reading 31] 2 [32 32 Oxford 32] 5 [33 33 Sou
thend 33] 0 [34 34 Sheffield United 34] 4 [35 35 Birmingham 35]
] 4 [36 36 Stoke 36] 2 [37 37 Manchester City 37] 1 [38 38 Sw
indon 38] 1 [39 39 Port Vale 39] 1 [40 40 Wolverhampton 40] 1
1 Bradford 0 Played Friday : [41 41 Portsmouth 41] 1 Queen 's
Park Rangers 2 [42 42 Tranmere 42] 3 [43 43 Grimsby 43] 2
Playing Sunday : [44 44 Barnsley 44] v [45 45 Huddersfield 45]
Division two [46 46 Brentford 46] 3 [47 47 Luton 47] 2 [48 48
Bristol City 48] v [49 49 Blackpool 49] late kickoff [50 50 Bu
rnley 50] 2 [51 51 Walsall 51] 1 [52 52 Chesterfield 52] 1 [53
53 Bury 53] 2 [54 54 Peterborough 54] 2 [55 55 Crewe 55] 2 [56
56 Preston 56] 0 [57 57 Bristol Rovers 57] 0 [58 58 Rotherham
58] 1 [59 59 Shrewsbury 59] 2 [60 60 Stockport 60] 0 [61 61 No
tts County 61] 0 [62 62 Watford 62] 0 [63 63 Millwall 63] 2 [6
4 64 Wrexham 64] 4 2 Plymouth 4 [65 65 Wycombe 65] 1 [66 66 Gi
llingham 66] 1 3 York 1 [67 67 Bournemouth 67] 2 Division
three [68 68 Barnet 68] 1 [69 69 Wigan 69] 1 [70 70 Cardiff 70]
] 1 [71 71 Brighton 71] 0 [72 72 Carlisle 72] 0 [73 73 Hull 73]
] 0 [74 74 Chester 74] 1 4 Cambridge 1 [75 75 Darlington 75]
4 [76 76 Swansea 76] 1 5 Exeter 2 [77 77 Scarborough 77] 2 [78
78 Hartlepool 78] 2 [79 79 Fulham 79] 1 [80 80 Hereford 80] 1
[81 81 Doncaster 81] 0 6 Lincoln 1 [82 82 Leyton Orient 82] 1
[83 83 Northampton 83] 3 [84 84 Mansfield 84] 0 [85 85 Rochda
le 85] 0 [86 86 Colchester 86] 0 [87 87 Scunthorpe 87] 1 [88 8
8 Torquay 88] 0 Add Division two [89 89 Bristol City 89] 0 [90
90 Blackpool 90] 1
1: 9316 England
2: 17867 London
3: 9316 England
4: 42173 Aston_Villa_F.C.
5: 4721303 Derby_County_F.C.
6: 7473 Chelsea_F.C.
7: 435021 Coventry_City_F.C.
8: 298602 Leicester_City_F.C.
9: 2174 Arsenal_F.C.
10: 18119 Liverpool_F.C.
11: 184474 Sunderland_A.F.C.
12: 57802 Newcastle_United_F.C.

13: 177538 Sheffield_Wednesday_F.C.
14: 157294 Nottingham_Forest_F.C.
15: 298603 Middlesbrough_F.C.
16: 68198 Tottenham_Hotspur_F.C.
17: 91155 Everton_F.C.
18: 46417 West_Ham_United_F.C.
19: 81576 Southampton_F.C.
20: 19961 Manchester_United_F.C.
21: 298599 Blackburn_Rovers_F.C.
22: 7609513 Leeds_United_A.F.C.
23: 740757 Wimbledon_F.C.
24: 298601 Bolton_Wanderers_F.C.
25: 248083 Norwich_City_F.C.
26: 6720 Charlton_Athletic_F.C.
27: 33921 West_Bromwich_Albion_F.C.
28: 385313 Crystal_Palace_F.C.
29: 451131 Oldham_Athletic_A.F.C.
30: 352843 Ipswich_Town_F.C.
31: 168007 Reading_F.C.
32: 451165 Oxford_United_F.C.
33: 451168 Southend_United_F.C.
34: 314933 Sheffield_United_F.C.
35: 185196 Birmingham_City_F.C.
36: 203434 Stoke_City_F.C.
37: 165813 Manchester_City_F.C.
38: 74291 Swindon_Town_F.C.
39: 203456 Port_Vale_F.C.
40: 217730 Wolverhampton_Wanderers_F.C.
41: 564301 Portsmouth_F.C.
42: 69219 Tranmere_Rovers_F.C.
43: 451129 Grimsby_Town_F.C.
44: 333188 Barnsley_F.C.
45: 236211 Huddersfield_Town_F.C.
46: 451122 Brentford_F.C.
47: 18367 Luton_Town_F.C.
48: 4392 Bristol_City_F.C.
49: 245357 Blackpool_F.C.
50: 376725 Burnley_F.C.
51: 359501 Walsall_F.C.
52: 451126 Chesterfield_F.C.
53: 451156 Bury_F.C.
54: 451132 Peterborough_United_F.C.
55: 69438 Crewe_Alexandra_F.C.
56: 429092 Preston_North_End_F.C.
57: 241161 Bristol_Rovers_F.C.
58: 386362 Rotherham_United_F.C.
59: 451207 Shrewsbury_Town_F.C.
60: 451135 Stockport_County_F.C.
61: 448599 Notts_County_F.C.
62: 85099 Watford_F.C.
63: 435024 Millwall_F.C.
64: 413286 Wrexham_F.C.

65: 451136 Wycombe_Wanderers_F.C.
66: 165278 Gillingham_F.C.
67: 451121 A.F.C._Bournemouth
68: 451173 Barnet_F.C.
69: 367030 Wigan_Athletic_F.C.
70: 5624949 Cardiff_City_F.C.
71: 69455 Brighton_&_Hove_Albiaon_F.C.
72: 451157 Carlisle_United_F.C.
73: 451358 Hull_City_A.F.C.
74: 451195 Chester_City_F.C.
75: 451159 Darlington_F.C.
76: 451169 Swansea_City_A.F.C.
77: 946449 Scarborough_F.C.
78: 451130 Hartlepool_United_F.C.
79: 11228 Fulham_F.C.
80: 451201 Hereford_United_F.C.
81: 451160 Doncaster_Rovers_F.C.
82: 451162 Leyton_Orient_F.C.
83: 425750 Northampton_Town_F.C.
84: 451164 Mansfield_Town_F.C.
85: 451166 Rochdale_A.F.C.
86: 451127 Colchester_United_F.C.
87: 451167 Scunthorpe_United_F.C.
88: 193049 Torquay_United_F.C.
89: 4392 Bristol_City_F.C.
90: 245357 Blackpool_F.C.
1: 431144 Bradford_City_A.F.C.
[winner: 1095011 Bradford_Bulls, pem_pos=9,
final_scores=-0.3648771, final_local_scores=-0.42995375,
pem=-3.818427, similarity=12.024236]
RECALL MISS: 431144 Bradford_City_A.F.C.]
2: 23857 Plymouth_Argyle_F.C.
[winner: 23508 Plymouth, pem_pos=0,
final_scores=-0.60506564, final_local_scores=-0.6355895,
pem=-0.45987546, similarity=-12.376671]
RECALL MISS: 23857 Plymouth_Argyle_F.C.]
3: 451170 York_City_F.C.
[winner: 992980 York_City_Knights, pem_pos=18,
final_scores=-0.6988038, final_local_scores=-0.77701765,
pem=-4.2129073, similarity=3.8124006]
RECALL MISS: 451170 York_City_F.C.]
4: 243970 Cambridge_United_F.C.
[winner: 25978572 University_of_Cambridge, pem_pos=1,
final_scores=-0.87068945, final_local_scores=-0.8950469,
pem=-1.7952901, similarity=-12.8572035]
RECALL MISS: 243970 Cambridge_United_F.C.]
5: 347780 Exeter_City_F.C.
[winner: 30871776 Exeter_Chiefs, pem_pos=19,
final_scores=-0.39981967, final_local_scores=-0.43902627,
pem=-3.6095665, similarity=10.617677]
RECALL MISS: 347780 Exeter_City_F.C.]
6: 451163 Lincoln_City_F.C.

```
[winner: 62705 Lincoln_Motor_Company, pem_pos=2,
final_scores=-0.65906125, final_local_scores=-0.6117686,
pem=-2.4392905, similarity=-0.9014417]
RECALL MISS: 451163 Lincoln_City_F.C.]
1: 9316 England
    [pem_pos=1, final_scores=0.51037145,
final_local_scores=0.5775263, pem=-1.3527626,
similarity=28.68078]
2: 17867 London
    [pem_pos=0, final_scores=0.6271986,
final_local_scores=0.7180898, pem=-0.09086737,
similarity=26.013096]
3: 9316 England
    [pem_pos=1, final_scores=0.6401798,
final_local_scores=0.71932214, pem=-1.3527626,
similarity=32.912254]
4: 42173 Aston_Villa_F.C.
    [pem_pos=0, final_scores=0.4446252,
final_local_scores=0.37310034, pem=0.0, similarity=15.223771]
5: 4721303 Derby_County_F.C.
    [pem_pos=5, final_scores=0.8662794,
final_local_scores=0.8295403, pem=-3.703934,
similarity=48.98742]
6: 7473 Chelsea_F.C.
    [pem_pos=0, final_scores=0.72805566,
final_local_scores=0.70350355, pem=-0.51707125,
similarity=27.895575]
7: 435021 Coventry_City_F.C.
    [pem_pos=2, final_scores=0.85731834,
final_local_scores=0.81840855, pem=-3.1334689,
similarity=45.552948]
8: 298602 Leicester_City_F.C.
    [pem_pos=3, final_scores=0.82396907,
final_local_scores=0.7845938, pem=-3.2133117,
similarity=44.978046]
9: 2174 Arsenal_F.C.
    [pem_pos=0, final_scores=0.5277857,
final_local_scores=0.47478443, pem=-0.15820743,
similarity=19.118584]
10: 18119 Liverpool_F.C.
    [pem_pos=1, final_scores=0.7229261,
final_local_scores=0.68963486, pem=-1.1702383,
similarity=31.033731]
11: 184474 Sunderland_A.F.C.
    [pem_pos=0, final_scores=0.75858027,
final_local_scores=0.71402866, pem=-0.39640337,
similarity=27.553455]
12: 57802 Newcastle_United_F.C.
    [pem_pos=2, final_scores=0.82710487,
final_local_scores=0.7906208, pem=-2.4885974,
similarity=41.216793]
13: 177538 Sheffield_Wednesday_F.C.
```

```
[pem_pos=0, final_scores=0.3007638,
final_local_scores=0.23596048, pem=0.0, similarity=11.131245]
14: 157294 Nottingham_Forest_F.C.
    [pem_pos=0, final_scores=0.49291754,
final_local_scores=0.42251158, pem=0.0, similarity=16.698301]
15: 298603 Middlesbrough_F.C.
    [pem_pos=0, final_scores=0.6112526,
final_local_scores=0.55356276, pem=-0.5175458,
similarity=23.423624]
16: 68198 Tottenham_Hotspur_F.C.
    [pem_pos=1, final_scores=0.46070576,
final_local_scores=0.3964768, pem=-0.99701256,
similarity=21.34328]
17: 91155 Everton_F.C.
    [pem_pos=0, final_scores=0.5323785,
final_local_scores=0.47118306, pem=-0.17171584,
similarity=19.084574]
18: 46417 West_Ham_United_F.C.
    [pem_pos=0, final_scores=0.5957286,
final_local_scores=0.53926986, pem=-0.5127755,
similarity=22.971155]
19: 81576 Southampton_F.C.
    [pem_pos=1, final_scores=0.77195805,
final_local_scores=0.73795706, pem=-1.1503892,
similarity=32.367817]
20: 19961 Manchester_United_F.C.
    [pem_pos=0, final_scores=0.39509434,
final_local_scores=0.3347267, pem=0.0, similarity=14.078625]
21: 298599 Blackburn_Rovers_F.C.
    [pem_pos=1, final_scores=0.7085882,
final_local_scores=0.65799767, pem=-2.734066,
similarity=38.593952]
22: 7609513 Leeds_United_A.F.C.
    [pem_pos=4, final_scores=0.74103445,
final_local_scores=0.6960005, pem=-3.421221,
similarity=43.46489]
23: 740757 Wimbledon_F.C.
    [pem_pos=1, final_scores=0.68750006,
final_local_scores=0.6553096, pem=-1.8899087,
similarity=33.923077]
24: 298601 Bolton_Wanderers_F.C.
    [pem_pos=2, final_scores=0.5980806,
final_local_scores=0.5365806, pem=-2.2845325,
similarity=32.525993]
25: 248083 Norwich_City_F.C.
    [pem_pos=2, final_scores=0.94531673,
final_local_scores=0.9156979, pem=-2.7278075,
similarity=46.250206]
26: 6720 Charlton_Athletic_F.C.
    [pem_pos=1, final_scores=0.8008217,
final_local_scores=0.7629022, pem=-2.103781,
similarity=38.296925]
```

27: 33921 West_Bromwich_Albion_F.C.
[pem_pos=2, final_scores=0.62791055,
final_local_scores=0.5720698, pem=-1.5002033,
similarity=29.319756]
28: 385313 Crystal_Palace_F.C.
[pem_pos=0, final_scores=0.5301052,
final_local_scores=0.46370035, pem=-0.51090026,
similarity=20.705812]
29: 451131 Oldham_Athletic_A.F.C.
[pem_pos=3, final_scores=0.95550066,
final_local_scores=0.92960244, pem=-2.8632221,
similarity=47.40155]
30: 352843 Ipswich_Town_F.C.
[pem_pos=4, final_scores=0.86127,
final_local_scores=0.8283732, pem=-2.6931517,
similarity=43.4558]
31: 168007 Reading_F.C.
[pem_pos=0, final_scores=0.8987923,
final_local_scores=0.86643285, pem=-1.2007197,
similarity=36.4755]
32: 451165 Oxford_United_F.C.
[pem_pos=18, final_scores=0.99933606,
final_local_scores=0.9831651, pem=-4.3872933,
similarity=57.288105]
33: 451168 Southend_United_F.C.
[pem_pos=2, final_scores=0.75154334,
final_local_scores=0.70765704, pem=-1.771178,
similarity=34.83955]
34: 314933 Sheffield_United_F.C.
[pem_pos=0, final_scores=0.34399652,
final_local_scores=0.25918567, pem=0.0, similarity=11.82433]
35: 185196 Birmingham_City_F.C.
[pem_pos=2, final_scores=0.66508836,
final_local_scores=0.61185795, pem=-3.102209,
similarity=39.21907]
36: 203434 Stoke_City_F.C.
[pem_pos=0, final_scores=0.5764307,
final_local_scores=0.5186927, pem=-0.3417574,
similarity=21.427067]
37: 165813 Manchester_City_F.C.
[pem_pos=0, final_scores=0.3088733,
final_local_scores=0.23169586, pem=0.0, similarity=11.00398]
38: 74291 Swindon_Town_F.C.
[pem_pos=4, final_scores=0.7707179,
final_local_scores=0.7295018, pem=-2.489671,
similarity=39.39872]
39: 203456 Port_Vale_F.C.
[pem_pos=0, final_scores=0.56532604,
final_local_scores=0.51548284, pem=0.0, similarity=19.47275]
40: 217730 Wolverhampton_Wanderers_F.C.
[pem_pos=1, final_scores=0.90132564,
final_local_scores=0.86601716, pem=-2.274954,

```
similarity=42.304947]
41: 564301 Portsmouth_F.C.
    [pem_pos=1, final_scores=0.77400583,
final_local_scores=0.73010355, pem=-1.2302613,
similarity=32.56781]
42: 69219 Tranmere_Rovers_F.C.
    [pem_pos=0, final_scores=0.48716044,
final_local_scores=0.42408162, pem=-0.6002322,
similarity=20.009312]
43: 451129 Grimsby_Town_F.C.
    [pem_pos=2, final_scores=1.0024871,
final_local_scores=0.9813824, pem=-2.6069987,
similarity=47.553387]
44: 333188 Barnsley_F.C.
    [pem_pos=0, final_scores=0.41903096,
final_local_scores=0.34181908, pem=-0.7126868,
similarity=18.165977]
45: 236211 Huddersfield_Town_F.C.
    [pem_pos=3, final_scores=0.7570774,
final_local_scores=0.71191674, pem=-2.2859883,
similarity=37.76629]
46: 451122 Brentford_F.C.
    [pem_pos=0, final_scores=0.6871598,
final_local_scores=0.6383555, pem=-0.36028567,
similarity=25.098804]
47: 18367 Luton_Town_F.C.
    [pem_pos=2, final_scores=0.6626616,
final_local_scores=0.60832506, pem=-2.5537767,
similarity=36.13118]
48: 4392 Bristol_City_F.C.
    [pem_pos=0, final_scores=0.46648312,
final_local_scores=0.39738727, pem=0.0, similarity=15.948542]
49: 245357 Blackpool_F.C.
    [pem_pos=1, final_scores=0.74045587,
final_local_scores=0.6928299, pem=-0.8537509,
similarity=29.407969]
50: 376725 Burnley_F.C.
    [pem_pos=1, final_scores=0.31460822,
final_local_scores=0.22987026, pem=-0.9914584,
similarity=16.341206]
51: 359501 Walsall_F.C.
    [pem_pos=1, final_scores=0.50797665,
final_local_scores=0.44164795, pem=-0.61973774,
similarity=20.639599]
52: 451126 Chesterfield_F.C.
    [pem_pos=1, final_scores=0.9081617,
final_local_scores=0.88300806, pem=-1.2678006,
similarity=37.33493]
53: 451156 Bury_F.C.
    [pem_pos=0, final_scores=0.83158165,
final_local_scores=0.79308486, pem=-0.72067994,
similarity=31.676117]
```

54: 451132 Peterborough_United_F.C.
[pem_pos=6, final_scores=0.880658,
final_local_scores=0.8502462, pem=-3.1818013,
similarity=46.765884]
55: 69438 Crewe_Alexandra_F.C.
[pem_pos=3, final_scores=0.72576606,
final_local_scores=0.6824335, pem=-2.8374243,
similarity=39.885242]
56: 429092 Preston_North_End_F.C.
[pem_pos=4, final_scores=0.7826982,
final_local_scores=0.73783904, pem=-3.194944,
similarity=43.482903]
57: 241161 Bristol_Rovers_F.C.
[pem_pos=0, final_scores=0.43568867,
final_local_scores=0.36448896, pem=0.0, similarity=14.966791]
58: 386362 Rotherham_United_F.C.
[pem_pos=3, final_scores=0.53137326,
final_local_scores=0.4684391, pem=-2.6427798,
similarity=32.440716]
59: 451207 Shrewsbury_Town_F.C.
[pem_pos=12, final_scores=1.0103474,
final_local_scores=0.99370664, pem=-3.2502756,
similarity=51.419407]
60: 451135 Stockport_County_F.C.
[pem_pos=4, final_scores=0.6622483,
final_local_scores=0.6161459, pem=-2.5287716,
similarity=36.228592]
61: 448599 Notts_County_F.C.
[pem_pos=0, final_scores=0.4337042,
final_local_scores=0.36068004, pem=0.0, similarity=14.853125]
62: 85099 Watford_F.C.
[pem_pos=1, final_scores=0.7677868,
final_local_scores=0.72407347, pem=-0.95984375,
similarity=30.91729]
63: 435024 Millwall_F.C.
[pem_pos=0, final_scores=0.5405982,
final_local_scores=0.47603536, pem=0.0, similarity=18.295557]
64: 413286 Wrexham_F.C.
[pem_pos=1, final_scores=0.529517,
final_local_scores=0.47166312, pem=-0.88259363,
similarity=22.964762]
65: 451136 Wycombe_Wanderers_F.C.
[pem_pos=1, final_scores=0.6815321,
final_local_scores=0.6410794, pem=-1.147249,
similarity=29.45972]
66: 165278 Gillingham_F.C.
[pem_pos=0, final_scores=0.6427645,
final_local_scores=0.5911202, pem=-0.2568813,
similarity=23.126879]
67: 451121 A.F.C._Bournemouth
[pem_pos=1, final_scores=0.57485723,
final_local_scores=0.5137387, pem=-1.0576377,

```
similarity=25.172298]
68: 451173 Barnet_F.C.
    [pem_pos=0, final_scores=0.7097843,
final_local_scores=0.67661196, pem=-0.6797195,
similarity=27.977583]
69: 367030 Wigan_Athletic_F.C.
    [pem_pos=2, final_scores=0.53605634,
final_local_scores=0.4758188, pem=-2.32737,
similarity=30.945694]
70: 5624949 Cardiff_City_F.C.
    [pem_pos=3, final_scores=0.8329659,
final_local_scores=0.7945476, pem=-3.4864018,
similarity=46.760193]
71: 69455 Brighton_&_Hove_Albion_F.C.
    [pem_pos=3, final_scores=0.98698515,
final_local_scores=0.9656618, pem=-3.1271749,
similarity=49.91305]
72: 451157 Carlisle_United_F.C.
    [pem_pos=12, final_scores=0.9747651,
final_local_scores=0.95333296, pem=-3.5999887,
similarity=52.116367]
73: 451358 Hull_City_A.F.C.
    [pem_pos=11, final_scores=0.84352285,
final_local_scores=0.8066805, pem=-3.7682815,
similarity=48.65517]
74: 451195 Chester_City_F.C.
    [pem_pos=1, final_scores=0.97022563,
final_local_scores=0.96164316, pem=-3.2160804,
similarity=50.276608]
75: 451159 Darlington_F.C.
    [pem_pos=1, final_scores=0.87641746,
final_local_scores=0.847079, pem=-1.8695662,
similarity=39.53523]
76: 451169 Swansea_City_A.F.C.
    [pem_pos=3, final_scores=0.97512275,
final_local_scores=0.955417, pem=-2.5988693,
similarity=46.734318]
77: 946449 Scarborough_F.C.
    [pem_pos=2, final_scores=0.9255678,
final_local_scores=0.91683024, pem=-1.8160111,
similarity=41.325504]
78: 451130 Hartlepool_United_F.C.
    [pem_pos=1, final_scores=0.81303376,
final_local_scores=0.77670175, pem=-1.75166,
similarity=36.793842]
79: 11228 Fulham_F.C.
    [pem_pos=0, final_scores=0.80139035,
final_local_scores=0.7671081, pem=-0.3209339,
similarity=28.727037]
80: 451201 Hereford_United_F.C.
    [pem_pos=6, final_scores=1.2000537,
final_local_scores=1.2068894, pem=-2.7578042,
```

```

similarity=55.103065]
81: 451160 Doncaster_Rovers_F.C.
    [pem_pos=6, final_scores=0.88912207,
final_local_scores=0.85982007, pem=-2.950685,
similarity=45.794743]
82: 451162 Leyton_Orient_F.C.
    [pem_pos=0, final_scores=0.45701152,
final_local_scores=0.38679922, pem=0.0, similarity=15.632573]
83: 425750 Northampton_Town_F.C.
    [pem_pos=10, final_scores=0.96598226,
final_local_scores=0.94306666, pem=-3.2381344,
similarity=49.842182]
84: 451164 Mansfield_Town_F.C.
    [pem_pos=16, final_scores=1.0280094,
final_local_scores=1.0135155, pem=-4.267106,
similarity=57.540226]
85: 451166 Rochdale_A.F.C.
    [pem_pos=1, final_scores=0.97795707,
final_local_scores=0.9581154, pem=-0.95097107,
similarity=37.853317]
86: 451127 Colchester_United_F.C.
    [pem_pos=7, final_scores=1.0885373,
final_local_scores=1.0921257, pem=-2.8568413,
similarity=52.21687]
87: 451167 Scunthorpe_United_F.C.
    [pem_pos=1, final_scores=0.68641293,
final_local_scores=0.6383248, pem=-1.507785,
similarity=31.338167]
88: 193049 Torquay_United_F.C.
    [pem_pos=2, final_scores=0.66613847,
final_local_scores=0.61599094, pem=-2.019306,
similarity=33.45341]
89: 4392 Bristol_City_F.C.
    [pem_pos=0, final_scores=0.39406025,
final_local_scores=0.3182763, pem=0.0, similarity=13.587712]
90: 245357 Blackpool_F.C.
    [pem_pos=1, final_scores=0.8131483,
final_local_scores=0.77223545, pem=-0.8537509,
similarity=31.77759]

```

Sample: b'255_CRICKET'

CRICKET - [1 1 PAKISTAN 1] 318-2 V [2 2 ENGLAND 2] --
LUNCH . [3 3 LONDON 3] 1996-08-24 [4 4 Pakistan 4] were 318
for two at lunch on the third day of the third and final test
at [5 5 The Oval 5] on Saturday in reply to [6 6 England 6]
's 326 all out . Score : [7 7 England 7] 326 (J. Crawley
106 , G. Thorpe 54 . [8 8 Waqar Younis 8] 4-95) . [9 9 Pakis
tan 9] 318-2 ([10 10 Saeed Anwar 10] 169 not out , [11 11 Ij
az Ahmed 11] 61) .

1: 759295 Pakistan_national_cricket_team
2: 759125 England_cricket_team
3: 17867 London

```
4: 759295 Pakistan_national_cricket_team
5: 387687 The_Oval
6: 759125 England_cricket_team
7: 759125 England_cricket_team
8: 894433 Waqar_Younis
9: 759295 Pakistan_national_cricket_team
10: 894437 Saeed_Anwar
11: 938863 Ijaz_Ahmed_(cricketer)
1: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.77722114,
final_local_scores=0.7440341, pem=-2.7556365,
similarity=41.278755]
2: 759125 England_cricket_team
    [pem_pos=3, final_scores=0.65558,
final_local_scores=0.61046785, pem=-3.3054674,
similarity=40.282936]
3: 17867 London
    [pem_pos=0, final_scores=0.50898844,
final_local_scores=0.59778184, pem=-0.09086737,
similarity=22.422865]
4: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.8397892,
final_local_scores=0.81238025, pem=-2.7556365,
similarity=43.31834]
5: 387687 The_Oval
    [pem_pos=0, final_scores=0.61673415,
final_local_scores=0.61070865, pem=0.0, similarity=22.314478]
6: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.8229092,
final_local_scores=0.79324955, pem=-3.3054674,
similarity=45.737507]
7: 759125 England_cricket_team
    [pem_pos=3, final_scores=0.76451236,
final_local_scores=0.7294598, pem=-3.3054674,
similarity=43.833893]
8: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.6298366,
final_local_scores=0.640651, pem=0.0, similarity=23.208015]
9: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7415915,
final_local_scores=0.7051142, pem=-2.7556365,
similarity=40.117306]
10: 894437 Saeed_Anwar
    [pem_pos=0, final_scores=0.7888536,
final_local_scores=0.808362, pem=0.0, similarity=28.212847]
11: 938863 Ijaz_Ahmed_(cricketer)
    [pem_pos=0, final_scores=0.57358986,
final_local_scores=0.5853947, pem=0.0, similarity=21.559057]
```

Sample: b'256_TENNIS'
TENNIS - RESULTS AT [1 1 CANADIAN OPEN 1] . [2 2 TORONTO 2]
1996-08-23 Results from the [3 3 Canadian Open 3] tennis

tournament on Friday (prefix numbers denotes seedings) :
Quarterfinals 3 - [4 4 Wayne Ferreira 4] ([5 5 South Africa 5]
]) beat 5 - [6 6 Thomas Enqvist 6] ([7 7 Sweden 7]) 7-5
6-2 4 - [8 8 Marcelo Rios 8] ([9 9 Chile 9]) beat [10 10 Pa
trick Rafter 10] ([11 11 Australia 11]) 0-6 7-6 (7-4) 6-1
7 - [12 12 Todd Martin 12] ([13 13 U.S. 13]) beat Alex
O'Brien ([14 14 U.S. 14]) 6-4 6-4 [15 15 Todd Woodbridge 15]
([16 16 Australia 16]) beat [17 17 Mark Philippoussis 17] ([18 18 Australia 18]) 7-5 6-4

1: 1872219 Canadian_Open_(tennis)
2: 64646 Toronto
3: 1872219 Canadian_Open_(tennis)
4: 1723516 Wayne_Ferreira
5: 17416221 South_Africa
6: 933709 Thomas_Enqvist
7: 5058739 Sweden
8: 483554 Marcelo_Rios
9: 5489 Chile
10: 23297 Patrick_Rafter
11: 4689264 Australia
12: 871725 Todd_Martin
13: 3434750 United_States
14: 3434750 United_States
15: 1547951 Todd_Woodbridge
16: 4689264 Australia
17: 261368 Mark_Philippoussis
18: 4689264 Australia

1: 1872219 Canadian_Open_(tennis)
[pem_pos=0, final_scores=0.53222555,
final_local_scores=0.55831546, pem=-0.58982956,
similarity=23.958544]
2: 64646 Toronto
[pem_pos=0, final_scores=0.5673867,
final_local_scores=0.6105957, pem=-0.12291032,
similarity=22.979513]
3: 1872219 Canadian_Open_(tennis)
[pem_pos=0, final_scores=0.45040458,
final_local_scores=0.46893847, pem=-0.58982956,
similarity=21.291357]
4: 1723516 Wayne_Ferreira
[pem_pos=0, final_scores=0.79992193,
final_local_scores=0.8995312, pem=0.0, similarity=30.93352]
5: 17416221 South_Africa
[pem_pos=0, final_scores=0.5954992,
final_local_scores=0.6583769, pem=-0.26251382,
similarity=25.164583]
6: 933709 Thomas_Enqvist
[pem_pos=0, final_scores=0.8754726,
final_local_scores=0.9564118, pem=0.0, similarity=32.630947]
7: 5058739 Sweden
[pem_pos=0, final_scores=0.68747616,
final_local_scores=0.74723136, pem=-0.2135508,

```
similarity=27.54991]
8: 483554 Marcelo_Ríos
    [pem_pos=0, final_scores=0.70722055,
final_local_scores=0.7546752, pem=0.0, similarity=26.610727]
9: 5489 Chile
    [pem_pos=0, final_scores=0.6359982,
final_local_scores=0.6971519, pem=-0.27404457,
similarity=26.384413]
10: 23297 Patrick_Rafter
    [pem_pos=0, final_scores=0.9761959,
final_local_scores=1.052563, pem=0.0, similarity=35.50029]
11: 4689264 Australia
    [pem_pos=0, final_scores=0.7659028,
final_local_scores=0.808844, pem=-0.17168477,
similarity=29.16088]
12: 871725 Todd_Martin
    [pem_pos=0, final_scores=0.81532186,
final_local_scores=0.87893337, pem=0.0, similarity=30.318836]
13: 3434750 United_States
    [pem_pos=0, final_scores=0.46347886,
final_local_scores=0.48115343, pem=-0.14075686,
similarity=19.213747]
14: 3434750 United_States
    [pem_pos=0, final_scores=0.49269277,
final_local_scores=0.51306516, pem=-0.14075686,
similarity=20.166058]
15: 1547951 Todd_Woodbridge
    [pem_pos=0, final_scores=0.8425506,
final_local_scores=0.9339166, pem=0.0, similarity=31.959648]
16: 4689264 Australia
    [pem_pos=0, final_scores=0.6939451,
final_local_scores=0.7302412, pem=-0.17168477,
similarity=26.815216]
17: 261368 Mark_Philippoussis
    [pem_pos=0, final_scores=0.7921098,
final_local_scores=0.8434103, pem=0.0, similarity=29.258759]
18: 4689264 Australia
    [pem_pos=0, final_scores=0.6857408,
final_local_scores=0.7212792, pem=-0.17168477,
similarity=26.547771]
```

Sample: b'257_RUGBY'

[1 1 RUGBY UNION 1] - [2 2 NEW ZEALAND 2] WIN FIRST SERIES
IN [3 3 SOUTH AFRICA 3] . [4 4 PRETORIA 4] , [5 5 South Africa
5] 1996-08-24 [6 6 New Zealand 6] made history on Saturday
when they completed their first series victory in [7 7 South A
frica 7] with a 33-26 victory in the second test . The win
gave the [8 8 All Blacks 8] an unbeatable 2-0 lead in the
three-test series . Each side scored three tries and the [9 9
Springboks 9] outscored the [10 10 All Blacks 10] 15-12 in
the second half but still suffered a fourth successive defeat
against their old enemies . Two tries from wing [11 11 Jeff Wi

lson 11] in the first quarter gave [12 12 New Zealand 12] a 24-11 lead before tries from flanker [13 13 Ruben Kruger 13] and scrum-half [14 14 Joost van der Westhuizen 14] in the space of two minutes narrowed the gap to a single point at 23-24 . The [15 15 Springboks 15] would have gone ahead had not fly-half [16 16 Joel Stransky 16] 's conversion hit an upright and [17 17 New Zealand 17] only scrambled to safety through a fine penalty from replacement fly-half Jon Preston and a drop goal by number eight [18 18 Zinzan Brooke 18] .

```
1: 25405 Rugby_union
2: 154957 New_Zealand_national_rugby_union_team
3: 376115 South_Africa_national_rugby_union_team
4: 25002 Pretoria
5: 17416221 South_Africa
6: 154957 New_Zealand_national_rugby_union_team
7: 376115 South_Africa_national_rugby_union_team
8: 154957 New_Zealand_national_rugby_union_team
9: 376115 South_Africa_national_rugby_union_team
10: 154957 New_Zealand_national_rugby_union_team
11: 905152 Jeff_Wilson_(sportsman)
12: 154957 New_Zealand_national_rugby_union_team
13: 21904572 Ruben_Kruger
14: 1748767 Joost_van_der_Westhuizen
15: 376115 South_Africa_national_rugby_union_team
16: 690902 Joel_Stransky
17: 154957 New_Zealand_national_rugby_union_team
18: 740724 Zinzan_Brooke
1: 25405 Rugby_union
    [pem_pos=0, final_scores=0.49296874,
final_local_scores=0.48948425, pem=0.0, similarity=18.6969]
2: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=4, final_scores=0.578734,
final_local_scores=0.52210367, pem=-3.8192244,
similarity=40.43986]
3: 376115 South_Africa_national_rugby_union_team
    [pem_pos=3, final_scores=0.48139977,
final_local_scores=0.41899037, pem=-3.598493,
similarity=36.162388]
4: 25002 Pretoria
    [pem_pos=0, final_scores=0.45668572,
final_local_scores=0.48143595, pem=0.0, similarity=18.456722]
5: 17416221 South_Africa
    [pem_pos=0, final_scores=0.49833876,
final_local_scores=0.54662067, pem=-0.26251382,
similarity=21.829556]
6: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=4, final_scores=0.6214995,
final_local_scores=0.56881845, pem=-3.8192244,
similarity=41.833923]
7: 376115 South_Africa_national_rugby_union_team
    [pem_pos=3, final_scores=0.46914762,
final_local_scores=0.40560675, pem=-3.598493,
```

```

similarity=35.762993]
8: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=0, final_scores=0.5648083,
final_local_scores=0.50689197, pem=0.0, similarity=19.216381]
9: 376115 South_Africa_national_rugby_union_team
    [pem_pos=0, final_scores=0.46905142,
final_local_scores=0.40550166, pem=0.0, similarity=16.190691]
10: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=0, final_scores=0.4991277,
final_local_scores=0.43514585, pem=0.0, similarity=17.075333]
11: 905152 Jeff_Wilson_(sportsman)
    [pem_pos=0, final_scores=0.7084822,
final_local_scores=0.7481119, pem=0.0, similarity=26.414864]
12: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=4, final_scores=0.63325423,
final_local_scores=0.5816587, pem=-3.8192244,
similarity=42.217106]
13: 21904572 Ruben_Kruger
    [pem_pos=0, final_scores=0.6774071,
final_local_scores=0.7370988, pem=0.0, similarity=26.086212]
14: 1748767 Joost_van_der_Westhuizen
    [pem_pos=0, final_scores=0.627461,
final_local_scores=0.6878565, pem=0.0, similarity=24.616722]
15: 376115 South_Africa_national_rugby_union_team
    [pem_pos=0, final_scores=0.5150869,
final_local_scores=0.45578843, pem=0.0, similarity=17.691349]
16: 690902 Joel_Stransky
    [pem_pos=0, final_scores=0.5880283,
final_local_scores=0.636921, pem=0.0, similarity=23.096706]
17: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=4, final_scores=0.6921194,
final_local_scores=0.64596, pem=-3.8192244,
similarity=44.135983]
18: 740724 Zinza_Brooke
    [pem_pos=0, final_scores=0.59612817,
final_local_scores=0.6470579, pem=0.0, similarity=23.399212]

```

Sample: b'258_RUGBY'

```

[1 1 RUGBY UNION 1] - [2 2 NEW ZEALAND 2] DEFEAT [3 3 SOUTH AFRICA 3] 33-26 . [4 4 PRETORIA 4] , [5 5 South Africa 5] 1996-08-24 [6 6 New Zealand 6] beat [7 7 South Africa 7] 33-26 ( halftime 21-11 ) in the second test on Saturday . Scorers : [8 8 South Africa 8] - Tries : [9 9 Hannes Strydom 9] , [10 10 Ruben Kruger 10] , [11 11 Joost van der Westhuizen 11] . Penalties : Joel Stranksy ( 3 ) . Conversion : Stranksy . [12 12 New Zealand 12] - Tries : [13 13 Jeff Wilson 13] ( 2 ) , [14 14 Zinza Brooke 14] . Penalties : [15 15 Simon Culhane 15] , Jon Preston ( 2 ) . Conversions : [16 16 Culhane 16] ( 3 ) . Drop goal : [17 17 Zinza Brooke 17] . [18 18 New Zealand 18] lead the three-test series 2-0 .
1: 25405 Rugby_union
2: 154957 New_Zealand_national_rugby_union_team

```

```
3: 376115 South_Africa_national_rugby_union_team
4: 25002 Pretoria
5: 17416221 South_Africa
6: 154957 New_Zealand_national_rugby_union_team
7: 376115 South_Africa_national_rugby_union_team
8: 376115 South_Africa_national_rugby_union_team
9: 21991758 Hannes_Strydom
10: 21904572 Ruben_Kruger
11: 1748767 Joost_van_der_Westhuizen
12: 154957 New_Zealand_national_rugby_union_team
13: 905152 Jeff_Wilson_(sportsman)
14: 740724 Zinzan_Brooke
15: 5210833 Simon_Culhane
16: 5210833 Simon_Culhane
17: 740724 Zinzan_Brooke
18: 154957 New_Zealand_national_rugby_union_team
1: 25405 Rugby_union
    [pem_pos=0, final_scores=0.48749685,
final_local_scores=0.48268455, pem=0.0, similarity=18.493982]
2: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=4, final_scores=0.5933524,
final_local_scores=0.5465584, pem=-3.8192244,
similarity=41.16964]
3: 376115 South_Africa_national_rugby_union_team
    [pem_pos=3, final_scores=0.52872545,
final_local_scores=0.4795326, pem=-3.598493,
similarity=37.969086]
4: 25002 Pretoria
    [pem_pos=0, final_scores=0.45110315,
final_local_scores=0.481143, pem=0.0, similarity=18.447979]
5: 17416221 South_Africa
    [pem_pos=0, final_scores=0.5192973,
final_local_scores=0.5779294, pem=-0.26251382,
similarity=22.76387]
6: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=4, final_scores=0.6163191,
final_local_scores=0.5716461, pem=-3.8192244,
similarity=41.91831]
7: 376115 South_Africa_national_rugby_union_team
    [pem_pos=3, final_scores=0.5495034,
final_local_scores=0.50306594, pem=-3.598493,
similarity=38.671368]
8: 376115 South_Africa_national_rugby_union_team
    [pem_pos=3, final_scores=0.5288636,
final_local_scores=0.47968352, pem=-3.598493,
similarity=37.97359]
9: 21991758 Hannes_Strydom
    [pem_pos=0, final_scores=0.611877,
final_local_scores=0.700253, pem=0.0, similarity=24.98666]
10: 21904572 Ruben_Kruger
    [pem_pos=0, final_scores=0.7102062,
final_local_scores=0.7771367, pem=0.0, similarity=27.281021]
```

```

11: 1748767 Joost_van_der_Westhuizen
    [pem_pos=0, final_scores=0.46272868,
final_local_scores=0.51341766, pem=0.0, similarity=19.41112]
12: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=4, final_scores=0.6224238,
final_local_scores=0.57831454, pem=-3.8192244,
similarity=42.117306]
13: 905152 Jeff_Wilson_(sportsman)
    [pem_pos=0, final_scores=0.6144646,
final_local_scores=0.63786966, pem=0.0, similarity=23.125015]
14: 740724 Zinzan_Brooke
    [pem_pos=0, final_scores=0.81323,
final_local_scores=0.87170106, pem=0.0, similarity=30.10301]
15: 5210833 Simon_Culhane
    [pem_pos=0, final_scores=0.68293285,
final_local_scores=0.7384236, pem=0.0, similarity=26.125746]
16: 5210833 Simon_Culhane
    [pem_pos=0, final_scores=0.54990065,
final_local_scores=0.6073834, pem=0.0, similarity=22.215244]
17: 740724 Zinzan_Brooke
    [pem_pos=0, final_scores=0.84298426,
final_local_scores=0.9042031, pem=0.0, similarity=31.072939]
18: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=4, final_scores=0.6199258,
final_local_scores=0.57558584, pem=-3.8192244,
similarity=42.035877]

```

Sample: b'259_SOCCER'

SOCCER - [1 1 YUGOSLAV 1] LEAGUE RESULTS / STANDINGS . [2 2 B
 ELGRADE 2] 1996-08-24 Results of [3 3 Yugoslav 3] league
 soccer matches played on Saturday : Division A Cukaricki 0 [4
 4 Hajduk 4] 2 Becej 2 [5 5 Borac 5] 0 Mladost (L) 0 [6 6 Ze
 mun 6] 0 [7 7 Rad 7] 1 Buducnost (P) 0 Standings
 (tabulated under played , won , drawn , lost , goals for ,
 against , points) : Becej 3 2 1 0 5 1 7 [8 8 Partizan 8] 2 2
 0 0 6 2 6 [9 9 Vojvodina 9] 2 2 0 0 4 1 6 [10 10 Red Star 10]
 2 2 0 0 3 1 6 Mladost (L) 3 1 1 1 6 4 4 [11 11 Rad 11] 3 1
 1 1 2 2 4 [12 12 Hajduk 12] 3 1 0 2 3 3 3 Cukaricki 3 1 0 2 5
 6 3 [13 13 Buducnost 13] 3 1 0 2 3 5 3 [14 14 Zemun 14] 3 0 2
 1 2 3 2 [15 15 Proleter 15] 2 0 1 1 1 4 1 [16 16 Borac 16] 3
 0 0 3 0 8 0 Division B [17 17 Sutjeska 17] 3 [18 18 Sloboda 18]
] 2 Loznica 0 Obilic 1 [19 19 OFK Kikinda 19] 1 Radnicki
 (N) 0 [20 20 Spartak 20] 1 Buducnost (V) 2 [21 21 OFK Beo
 grad 21] 2 Mladost (BJ) 2 Standings : Obilic 3 3 0 0 8 1 9
 Loznica 3 2 0 1 7 3 6 [1 1 Sutjeska 1] 3 2 0 1 6 3 6 [22 22 OF
 K Kikinda 22] 3 2 0 1 4 1 6 Buducnost (V) 3 2 0 1 4 3 6 [23
 23 Spartak 23] 3 1 1 1 3 3 4 Zeleznik 2 1 0 1 4 4 3 [24 24 OFK
 Beograd 24] 3 0 3 1 4 4 3 Radnicki 3 1 0 2 5 6 3 [25 25 Slobo
 da 25] 3 0 1 2 4 8 1 Mladost (BJ) 3 0 1 2 2 6 1 [26 26 Rudar
 26] 2 0 0 2 0 6 0

1: 34244 Yugoslavia
 2: 55904 Belgrade

3: 34244 Yugoslavia
4: 12587188 FK_Hajduk_Beograd
5: 4827027 FK_Borac_Čačak
6: 4826904 FK_Zemun
7: 1132478 FK_Rad
8: 381658 FK_Partizan
9: 2652716 FK_Vojvodina
10: 361966 Red_Star_Belgrade
11: 1132478 FK_Rad
12: 12587188 FK_Hajduk_Beograd
13: 2652690 FK_Budućnost_Podgorica
14: 4826904 FK_Zemun
15: 12576085 FK_Proleter_Zrenjanin
16: 4827027 FK_Borac_Čačak
17: 5769234 FK_Sutjeska_Nikšić
18: 1940216 FK_Sloboda_Tuzla
19: 17634183 FK_Kikinda
20: 20517361 FK_Spartak_Zlatibor_Voda
21: 2424381 OFK_Beograd
22: 17634183 FK_Kikinda
23: 20517361 FK_Spartak_Zlatibor_Voda
24: 2424381 OFK_Beograd
25: 1940216 FK_Sloboda_Tuzla
26: 5077936 NK_Rudar_Velenje
1: 12821345 FK_Sutjeska_Foča
 [winner: 5769234 FK_Sutjeska_Nikšić, pem_pos=0,
final_scores=0.08546887, final_local_scores=0.022092104,
pem=-0.5167451, similarity=7.559125]
 gt: 12821345 FK_Sutjeska_Foča, pem_pos=7,
final_scores=-0.018646121, final_local_scores=-0.017197192,
pem=-1.9231789, similarity=14.035061]
1: 34244 Yugoslavia
 [pem_pos=0, final_scores=0.26613206,
final_local_scores=0.28969866, pem=-0.7596219,
similarity=16.86584]
2: 55904 Belgrade
 [pem_pos=0, final_scores=0.44076365,
final_local_scores=0.4605214, pem=-0.05193398,
similarity=18.115015]
3: 34244 Yugoslavia
 [pem_pos=0, final_scores=0.26656824,
final_local_scores=0.29017514, pem=-0.7596219,
similarity=16.88006]
4: 12587188 FK_Hajduk_Beograd
 [pem_pos=5, final_scores=0.16230248,
final_local_scores=0.15149266, pem=-2.0794415,
similarity=19.918882]
5: 4827027 FK_Borac_Čačak
 [pem_pos=1, final_scores=0.27179253,
final_local_scores=0.2161577, pem=-0.9991784,
similarity=15.973978]
6: 4826904 FK_Zemun

```
[pem_pos=1, final_scores=0.47493774,
final_local_scores=0.43068057, pem=-0.8501668,
similarity=21.565418]
7: 1132478 FK_Rad
    [pem_pos=1, final_scores=0.6733965,
final_local_scores=0.651844, pem=-1.2102866,
similarity=30.123766]
8: 381658 FK_Partizan
    [pem_pos=0, final_scores=0.37411082,
final_local_scores=0.3134109, pem=-0.1934529,
similarity=14.494546]
9: 2652716 FK_Vojvodina
    [pem_pos=1, final_scores=0.3946004,
final_local_scores=0.33840132, pem=-1.2114366,
similarity=20.776264]
10: 361966 Red_Star_Belgrade
    [pem_pos=0, final_scores=0.40362722,
final_local_scores=0.34146583, pem=-1.0362908,
similarity=19.915245]
11: 1132478 FK_Rad
    [pem_pos=1, final_scores=0.6291774,
final_local_scores=0.6035413, pem=-1.2102866,
similarity=28.682316]
12: 12587188 FK_Hajduk_Beograd
    [pem_pos=5, final_scores=0.16446774,
final_local_scores=0.15385789, pem=-2.0794415,
similarity=19.989466]
13: 2652690 FK_Budućnost_Podgorica
    [pem_pos=1, final_scores=0.2253042,
final_local_scores=0.16789448, pem=-0.6931472,
similarity=12.869462]
14: 4826904 FK_Zemun
    [pem_pos=1, final_scores=0.49818188,
final_local_scores=0.45607132, pem=-0.8501668,
similarity=22.323132]
15: 12576085 FK_Proleter_Zrenjanin
    [pem_pos=0, final_scores=0.3410461,
final_local_scores=0.31500196, pem=0.0, similarity=13.49]
16: 4827027 FK_Borac_Čačak
    [pem_pos=1, final_scores=0.31253487,
final_local_scores=0.26066247, pem=-0.9991784,
similarity=17.302088]
17: 5769234 FK_Sutjeska_Nikšić
    [pem_pos=0, final_scores=0.22439192,
final_local_scores=0.17384446, pem=-0.5167451,
similarity=12.087718]
18: 1940216 FK_Sloboda_Tuzla
    [pem_pos=0, final_scores=0.36199087,
final_local_scores=0.33469522, pem=-0.42356145,
similarity=16.381079]
19: 17634183 FK_Kikinda
    [pem_pos=0, final_scores=0.43635547,
```

```
final_local_scores=0.4277916, pem=0.0, similarity=16.855867]
20: 20517361 FK_Spartak_Zlatibor_Voda
    [pem_pos=3, final_scores=0.27951568,
final_local_scores=0.21785742, pem=-1.8375496,
similarity=20.583893]
21: 2424381 OFK_Beograd
    [pem_pos=0, final_scores=0.53436726,
final_local_scores=0.4849438, pem=0.0, similarity=18.561403]
22: 17634183 FK_Kikinda
    [pem_pos=0, final_scores=0.39722568,
final_local_scores=0.38504827, pem=0.0, similarity=15.580321]
23: 20517361 FK_Spartak_Zlatibor_Voda
    [pem_pos=3, final_scores=0.25887436,
final_local_scores=0.19530994, pem=-1.8375496,
similarity=19.911032]
24: 2424381 OFK_Beograd
    [pem_pos=0, final_scores=0.5418902,
final_local_scores=0.4931615, pem=0.0, similarity=18.806635]
25: 1940216 FK_Sloboda_Tuzla
    [pem_pos=0, final_scores=0.3822385,
final_local_scores=0.35681272, pem=-0.42356145,
similarity=17.04111]
26: 5077936 NK_Rudar_Velenje
    [pem_pos=1, final_scores=0.05723895,
final_local_scores=0.029328004, pem=-1.1495787,
similarity=11.216507]
1: 5769234 FK_Sutjeska_Nikšić
    [pem_pos=0, final_scores=0.08546887,
final_local_scores=0.022092104, pem=-0.5167451,
similarity=7.559125]
```

Sample: b'260_SOCCER'

```
SOCCER - [1 1 POLISH 1] FIRST DIVISION RESULTS . [2 2 WARSAW
2] 1996-08-24 Results of [3 3 Polish 3] first division soccer
matches on Saturday : [4 4 Amica Wronki 4] 3 [5 5 Hutnik Krakow 5]
0 Sokol Tychy 5 [6 6 Lech Poznan 6] 3 [7 7 Rakow Czestochowa 7] 1
[8 8 Stomil Olsztyn 8] 4 [9 9 Wisla Krakow 9] 1 [10 10 Gornik Zabrze 10] 0
[11 11 Slask Wroclaw 11] 3 [12 12 Odra Wodzislaw 12] 1
[13 13 GKS Katowice 13] 1 [14 14 Polonia Warsaw 14] 0
[15 15 Zaglebie Lubin 15] 2 [16 16 LKS Lodz 16] 1 [17 17 Legia Warsaw 17] 3
[18 18 GKS Belchatow 18] 2
1: 22936 Poland
2: 32908 Warsaw
3: 22936 Poland
4: 677540 Amica_Wronki
5: 805862 Hutnik_Nowa_Huta
6: 677622 Lech_Poznań
7: 8538149 Raków_Częstochowa
8: 10410246 Stomil_Olsztyn_(football)
9: 677524 Wisła_Kraków
10: 677580 Górnik_Zabrze
11: 772146 Śląsk_Wrocław
```

12: 677572 Odra_Wodzisław
13: 677596 GKS_Katowice
14: 677603 Polonia_Warsaw
15: 677663 Zagłębie_Lubin
16: 677683 ŁKS_Łódź
17: 677536 Legia_Warsaw
18: 677651 GKS_Bełchatów
1: 22936 Poland
 [pem_pos=0, final_scores=0.636993,
final_local_scores=0.6926734, pem=-0.81518245,
similarity=29.193558]
2: 32908 Warsaw
 [pem_pos=0, final_scores=0.48337877,
final_local_scores=0.5470858, pem=-0.14300999,
similarity=21.193556]
3: 22936 Poland
 [pem_pos=0, final_scores=0.5635507,
final_local_scores=0.6124489, pem=-0.81518245,
similarity=26.7995]
4: 677540 Amica_Wronki
 [pem_pos=0, final_scores=0.6595941,
final_local_scores=0.67029715, pem=0.0, similarity=24.092716]
5: 805862 Hutnik_Nowa_Huta
 [pem_pos=0, final_scores=0.5277081,
final_local_scores=0.55891424, pem=0.0, similarity=20.76883]
6: 677622 Lech_Poznań
 [pem_pos=0, final_scores=0.6022736,
final_local_scores=0.58717984, pem=0.0, similarity=21.612331]
7: 8538149 Raków_Częstochowa
 [pem_pos=0, final_scores=0.55530405,
final_local_scores=0.5885582, pem=0.0, similarity=21.653463]
8: 10410246 Stomil_Olsztyn_(football)
 [pem_pos=0, final_scores=0.5200574,
final_local_scores=0.542671, pem=0.0, similarity=20.2841]
9: 677524 Wisła_Kraków
 [pem_pos=0, final_scores=0.6008285,
final_local_scores=0.56808424, pem=0.0, similarity=21.04248]
10: 677580 Górnik_Zabrze
 [pem_pos=0, final_scores=0.64537114,
final_local_scores=0.61750835, pem=0.0, similarity=22.517393]
11: 772146 Śląsk_Wrocław
 [pem_pos=0, final_scores=0.5209057,
final_local_scores=0.49738967, pem=0.0, similarity=18.932814]
12: 677572 Odra_Wodzisław
 [pem_pos=0, final_scores=0.63347834,
final_local_scores=0.6498869, pem=0.0, similarity=23.483635]
13: 677596 GKS_Katowice
 [pem_pos=0, final_scores=0.54236704,
final_local_scores=0.5253535, pem=0.0, similarity=19.76731]
14: 677603 Polonia_Warsaw
 [pem_pos=0, final_scores=0.57605857,
final_local_scores=0.54628617, pem=0.0, similarity=20.391981]

```
15: 677663 Zagłębie_Lubin
    [pem_pos=0, final_scores=0.624459,
final_local_scores=0.6116777, pem=0.0, similarity=22.343395]
16: 677683 ŁKS Łódź
    [pem_pos=0, final_scores=0.54752976,
final_local_scores=0.5166526, pem=0.0, similarity=19.507656]
17: 677536 Legia_Warsaw
    [pem_pos=0, final_scores=0.4966523,
final_local_scores=0.45823574, pem=0.0, similarity=17.764381]
18: 677651 GKS_Bełchatów
    [pem_pos=0, final_scores=0.543344,
final_local_scores=0.548461, pem=0.0, similarity=20.456884]
```

Sample: b'261_BASKETBALL'

BASKETBALL - [1 1 PHILIPPINE 1] PRO-LEAGUE RESULTS . MANILA
1996-08-24 Results of semi-final round games on Friday in the
[2 2 Philippine Basketball Association 2] second conference ,
which includes [3 3 American 3] players : [4 4 Alaska Milk 4]
beat [1 Purefoods Hotdogs 1] 103-95 (34-48 half-time) [5 5 G
inebra San Miguel 5] beat [6 6 Shell 6] 120-103 (65-56)

```
1: 23440 Philippines
2: 364158 Philippine_Basketball_Association
3: 3434750 United_States
4: 4047853 Alaska_Aces_(PBA)
5: 3647082 Barangay_Ginebra_San_Miguel
6: 4092880 Shell_Turbo_Chargers
1: 23440 Philippines
    [pem_pos=0, final_scores=0.47888273,
final_local_scores=0.55529994, pem=-0.5672473,
similarity=23.745749]
2: 364158 Philippine_Basketball_Association
    [pem_pos=0, final_scores=0.36265165,
final_local_scores=0.3363321, pem=0.0, similarity=14.126534]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.37847978,
final_local_scores=0.4857369, pem=-0.12994486,
similarity=19.291733]
4: 4047853 Alaska_Aces_(PBA)
    [pem_pos=0, final_scores=0.56349176,
final_local_scores=0.5505777, pem=0.0, similarity=20.52005]
5: 3647082 Barangay_Ginebra_San_Miguel
    [pem_pos=0, final_scores=0.3963322,
final_local_scores=0.3739164, pem=0.0, similarity=15.248124]
6: 4092880 Shell_Turbo_Chargers
    [pem_pos=4, final_scores=0.55808324,
final_local_scores=0.5712882, pem=-3.0383937,
similarity=37.66135]
1: 3044449 San_Mig_Super_Coffee_Mixers
    [pem_pos=0, final_scores=0.062400952,
final_local_scores=-0.003301695, pem=-0.027398957,
similarity=4.140184]
```

```

Sample: b'262_BASEBALL'
    BASEBALL - RESULTS OF [1 1 S. KOREAN 1] PRO-BASEBALL GAMES .
[2 2 SEOUL 2] 1996-08-24 Results of [3 3 South Korean 3] pro-
baseball games played on Friday . Samsung 13 Hyundai 3 Haitai
5 [4 4 Hanwha 4] 4 OB 4 Lotte 2 Ssangbangwool 1 LG 0
Standings after games played on Friday ( tabulate under won ,
drawn , lost , winning percentage , games behind first
place ) W D L PCT GB Haitai 62 2 40 .606 - Ssangbangwool 56 2
47 .543 6 1/2 [5 5 Hanwha 5] 55 1 47 .539 7 Hyundai 54 5 47 .
533 7 1/2 Samsung 47 5 53 .471 14 Lotte 43 5 52 .455 15 1/2
LG 44 5 56 .443 17 OB 40 5 59 .407 20 1/2
1: 27019 South_Korea
2: 19159283 Seoul
3: 27019 South_Korea
4: 4086558 Hanwha_Eagles
5: 4086558 Hanwha_Eagles
1: 27019 South_Korea
    [pem_pos=0, final_scores=0.55884665,
final_local_scores=0.5411177, pem=0.0, similarity=20.237745]
2: 19159283 Seoul
    [pem_pos=0, final_scores=0.49959213,
final_local_scores=0.49560285, pem=-0.0378014,
similarity=19.08506]
3: 27019 South_Korea
    [pem_pos=0, final_scores=0.52915996,
final_local_scores=0.5086896, pem=0.0, similarity=19.270025]
4: 4086558 Hanwha_Eagles
    [pem_pos=1, final_scores=0.5375663,
final_local_scores=0.567147, pem=-0.6931472,
similarity=24.783953]
5: 4086558 Hanwha_Eagles
    [pem_pos=1, final_scores=0.5702053,
final_local_scores=0.60280013, pem=-0.6931472,
similarity=25.847914]

```

```

Sample: b'263_BASEBALL'
    BASEBALL - [1 1 MAJOR LEAGUE 1] STANDINGS AFTER FRIDAY 'S
GAMES . [2 2 NEW YORK 2] 1996-08-24 [3 3 Major League Baseball
3] standings after games played on Friday ( tabulate under
won , lost , winning percentage and games behind ) : [4 4 AME
RICAN LEAGUE EASTERN DIVISION 4] W L PCT GB [5 5 NEW YORK 5]
73 54 .575 - [6 6 BALTIMORE 6] 67 60 .528 6 [7 7 BOSTON 7] 64
65 .496 10 [1 1 TORONTO 1] 60 69 .465 14 [8 8 DETROIT 8] 46
82 .359 27 1/2 [9 9 CENTRAL DIVISION 9] [10 10 CLEVELAND 10]
76 52 .594 - 2 CHICAGO 69 61 .531 8 [11 11 MINNESOTA 11] 64
64 .500 12 [12 12 MILWAUKEE 12] 61 68 .473 15 1/2 [13 13 KANS
AS CITY 13] 58 72 .446 19 [14 14 WESTERN DIVISION 14] [15 15 T
EXAS 15] 74 55 .574 - [16 16 SEATTLE 16] 66 61 .520 7 [17 17 O
AKLAND 17] 62 69 .473 13 3 CALIFORNIA 60 68 .469 13 1/2
SATURDAY , AUGUST 24TH SCHEDULE [18 18 SEATTLE 18] AT [19 19 B
OSTON 19] [20 20 MILWAUKEE 20] AT [21 21 CLEVELAND 21] 4 CALI
FORNIA AT [22 22 BALTIMORE 22] [2 5 TORONTO 2] AT 6 CHICAGO [2

```

3 23 OAKLAND 23] AT [3 7 NEW YORK 3] [24 24 DETROIT 24] AT [25
25 KANSAS CITY 25] [26 26 TEXAS 26] AT [27 27 MINNESOTA 27] [2
8 28 NATIONAL LEAGUE EASTERN DIVISION 28] W L PCT GB [29 29 AT
LANTA 29] 80 47 .630 - [30 30 MONTREAL 30] 69 58 .543 11 [31 3
1 NEW YORK 31] 59 70 .457 22 8 FLORIDA 59 70 .457 22 [32 32 PH
ILADELPHIA 32] 53 76 .411 28 [33 33 CENTRAL DIVISION 33] ST
LOUIS 68 60 .531 - [34 34 HOUSTON 34] 68 61 .527 1/2 [35 35 CI
NCINNATI 35] 64 63 .504 3 1/2 9 CHICAGO 63 63 .500 4 [36 36 PI
TTSBURGH 36] 55 73 .430 13 [37 37 WESTERN DIVISION 37] [38 38
SAN DIEGO 38] 70 60 .538 - 10 LOS ANGELES 68 60 .531 1 [39 39
COLORADO 39] 66 63 .512 3 1/2 [40 40 SAN FRANCISCO 40] 54 72 .
429 14 SATURDAY , AUGUST 24TH SCHEDULE 11 CHICAGO AT [41 41 AT
LANTA 41] ST LOUIS AT [42 42 HOUSTON 42] [43 43 NEW YORK 43]
AT [44 44 LOS ANGELES 44] [4 12 MONTREAL 4] AT [45 45 SAN FRA
NCISCO 45] [46 46 CINCINNATI 46] AT 13 FLORIDA [47 47 PITTSBU
RGH 47] AT [48 48 COLORADO 48] [49 49 PHILADELPHIA 49] AT [50
50 SAN DIEGO 50]

1: 38776 Major_League_Baseball
2: 645042 New_York_City
3: 38776 Major_League_Baseball
4: 2004582 American_League_East
5: 4848143 New_York_Yankees
6: 4340 Baltimore_Orioles
7: 4339 Boston_Red_Sox
8: 8579 Detroit_Tigers
9: 2456377 National_League_Central
10: 6652 Cleveland_Indians
11: 20050 Minnesota_Twins
12: 20152 Milwaukee_Brewers
13: 16883 Kansas_City_Royals
14: 2456474 National_League_West
15: 30857 Texas_Rangers_(baseball)
16: 27660 Seattle_Mariners
17: 22523 Oakland_Athletics
18: 27660 Seattle_Mariners
19: 4339 Boston_Red_Sox
20: 20152 Milwaukee_Brewers
21: 6652 Cleveland_Indians
22: 4340 Baltimore_Orioles
23: 22523 Oakland_Athletics
24: 8579 Detroit_Tigers
25: 16883 Kansas_City_Royals
26: 30857 Texas_Rangers_(baseball)
27: 20050 Minnesota_Twins
28: 2456294 National_League_East
29: 2140 Atlanta_Braves
30: 20153 Montreal_Expos
31: 4848143 New_York_Yankees
32: 23741 Philadelphia_Phillies
33: 2456377 National_League_Central
34: 13894 Houston_Astros
35: 6671 Cincinnati_Reds

```
36: 24270 Pittsburgh_Pirates
37: 2456474 National_League_West
38: 28417 San_Diego_Padres
39: 6669 Colorado_Rockies
40: 28416 San_Francisco_Giants
41: 2140 Atlanta_Braves
42: 13894 Houston_Astros
43: 4848143 New_York_Yankees
44: 18110 Los_Angeles
45: 28416 San_Francisco_Giants
46: 6671 Cincinnati_Reds
47: 24270 Pittsburgh_Pirates
48: 6669 Colorado_Rockies
49: 23741 Philadelphia_Phillies
50: 28417 San_Diego_Padres
1: 64646 Toronto
    [winner: 30847 Toronto_Blue_Jays, pem_pos=6,
final_scores=0.23179518, final_local_scores=0.14632392,
pem=-4.422637, similarity=32.507298|
     gt: 64646 Toronto, pem_pos=0, final_scores=0.059230432,
final_local_scores=0.08378117, pem=-0.12291032,
similarity=7.2583156 ]
2: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.7285585, final_local_scores=-0.8210389,
pem=-4.591283, similarity=4.556384|
     RECALL MISS: 5945 Chicago_White_Sox]
3: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 799305 California_State_Route_60, pem_pos=0,
final_scores=-0.5413827, final_local_scores=-0.468953,
pem=0.0, similarity=-9.90478|
     RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
4: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 7590031 California_Golden_Bears_football,
pem_pos=1, final_scores=-0.8353421,
final_local_scores=-0.9180979, pem=-4.4771376,
similarity=1.0392108|
     RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
5: 64646 Toronto
    [winner: 30847 Toronto_Blue_Jays, pem_pos=6,
final_scores=0.34165913, final_local_scores=0.26633358,
pem=-4.422637, similarity=36.088627|
     gt: 64646 Toronto, pem_pos=0, final_scores=0.13088258,
final_local_scores=0.16205025, pem=-0.12291032,
similarity=9.594021 ]
6: 5945 Chicago_White_Sox
    [winner: 6886 Chicago, pem_pos=0,
final_scores=-0.66541505, final_local_scores=-0.731064,
pem=-0.20658566, similarity=-16.603252|
     RECALL MISS: 5945 Chicago_White_Sox]
7: 645042 New_York_City
    [winner: 4848143 New_York_Yankees, pem_pos=11,
```

```
final_scores=0.14735784, final_local_scores=0.04796976,
pem=-5.0692368, similarity=33.088524]
    gt: 645042 New_York_City, pem_pos=1,
final_scores=0.09791063, final_local_scores=0.12993741,
pem=-1.7906687, similarity=17.705238 ]
8: 11278 Miami_Marlins
    [winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.72370124, final_local_scores=-0.79413027,
pem=-4.620281, similarity=5.5170884|
     RECALL MISS: 11278 Miami_Marlins]
9: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.76307356, final_local_scores=-0.85874134,
pem=-4.591283, similarity=3.4312685|
     RECALL MISS: 5945 Chicago_White_Sox]
10: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 18213 Los_Angeles_Dodgers, pem_pos=15,
final_scores=-0.13550228, final_local_scores=-0.26306504,
pem=-4.7524934, similarity=22.084126|
     RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
11: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.67418563, final_local_scores=-0.7616448,
pem=-4.591283, similarity=6.328823|
     RECALL MISS: 5945 Chicago_White_Sox]
12: 7954681 Montreal
    [winner: 20153 Montreal_Expos, pem_pos=8,
final_scores=0.382079, final_local_scores=0.3076812,
pem=-4.5098605, similarity=37.796856|
     gt: 7954681 Montreal, pem_pos=0,
final_scores=-0.031524613, final_local_scores=0.006086707,
pem=-0.11023275, similarity=4.8708153 ]
13: 11278 Miami_Marlins
    [winner: 18933066 Florida, pem_pos=0,
final_scores=-0.7060505, final_local_scores=-0.7331439,
pem=-0.15405534, similarity=-16.950989|
     RECALL MISS: 11278 Miami_Marlins]
1: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.58785325,
final_local_scores=0.5399099, pem=-0.47792417,
similarity=22.800726]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.37884867,
final_local_scores=0.43675756, pem=-1.7906687,
similarity=26.861364]
3: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.5413793,
final_local_scores=0.48914415, pem=0.0, similarity=18.68675]
4: 2004582 American_League_East
    [pem_pos=0, final_scores=0.47716838,
final_local_scores=0.47655284, pem=0.0, similarity=18.311]
5: 4848143 New_York_Yankees
```

```
[pem_pos=11, final_scores=0.5223654,
final_local_scores=0.45758682, pem=-5.0692368,
similarity=45.312313]
6: 4340 Baltimore_Orioles
[pem_pos=3, final_scores=1.3113968,
final_local_scores=1.3177502, pem=-4.046561,
similarity=65.41983]
7: 4339 Boston_Red_Sox
[pem_pos=5, final_scores=1.1537434,
final_local_scores=1.1488187, pem=-4.418931,
similarity=62.403587]
8: 8579 Detroit_Tigers
[pem_pos=4, final_scores=1.1800115,
final_local_scores=1.1745784, pem=-4.1401463,
similarity=61.656235]
9: 2456377 National_League_Central
[pem_pos=0, final_scores=0.57521087,
final_local_scores=0.5945552, pem=-0.98951703,
similarity=27.213573]
10: 6652 Cleveland_Indians
[pem_pos=12, final_scores=1.2728947,
final_local_scores=1.2747855, pem=-4.191953,
similarity=64.92835]
11: 20050 Minnesota_Twins
[pem_pos=9, final_scores=1.2118071,
final_local_scores=1.2114935, pem=-4.2906566,
similarity=63.57635]
12: 20152 Milwaukee_Brewers
[pem_pos=4, final_scores=0.5885629,
final_local_scores=0.53252965, pem=-3.40205,
similarity=38.482338]
13: 16883 Kansas_City_Royals
[pem_pos=5, final_scores=0.8260247,
final_local_scores=0.79199076, pem=-3.5915844,
similarity=47.25589]
14: 2456474 National_League_West
[pem_pos=1, final_scores=0.556058,
final_local_scores=0.5610173, pem=-1.2952092,
similarity=27.875135]
15: 30857 Texas_Rangers_(baseball)
[pem_pos=13, final_scores=1.2083812,
final_local_scores=1.2088745, pem=-4.928425,
similarity=66.966484]
16: 27660 Seattle_Mariners
[pem_pos=4, final_scores=0.7255481,
final_local_scores=0.6871665, pem=-3.8567436,
similarity=45.569702]
17: 22523 Oakland_Athletics
[pem_pos=7, final_scores=1.2639122,
final_local_scores=1.266635, pem=-3.9188845,
similarity=63.20013]
18: 27660 Seattle_Mariners
```

```
[pem_pos=4, final_scores=0.6817905,
final_local_scores=0.639368, pem=-3.8567436,
similarity=44.1433]
19: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.0173885,
final_local_scores=0.9998717, pem=-4.418931,
similarity=57.95871]
20: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.51641685,
final_local_scores=0.453721, pem=-3.40205,
similarity=36.130527]
21: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.2318319,
final_local_scores=1.2299306, pem=-4.191953,
similarity=63.58979]
22: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.0606736,
final_local_scores=1.0438731, pem=-4.046561,
similarity=57.246796]
23: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.2018597,
final_local_scores=1.1988521, pem=-3.9188845,
similarity=61.177353]
24: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.2573509,
final_local_scores=1.2590601, pem=-4.1401463,
similarity=64.17734]
25: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.828999,
final_local_scores=0.7952397, pem=-3.5915844,
similarity=47.352844]
26: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.2138464,
final_local_scores=1.2148445, pem=-4.928425,
similarity=67.14464]
27: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=0.93766445,
final_local_scores=0.9120342, pem=-4.2906566,
similarity=54.639893]
28: 2456294 National_League_East
    [pem_pos=0, final_scores=0.4342664,
final_local_scores=0.42880726, pem=0.0, similarity=16.886177]
29: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=1.1265874,
final_local_scores=1.1155562, pem=-4.136474,
similarity=59.874924]
30: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.8116686,
final_local_scores=0.7769314, pem=-4.5098605,
similarity=51.80022]
31: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.5598712,
```

```
final_local_scores=0.4985562, pem=-5.0692368,
similarity=46.534924]
32: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.2803831,
final_local_scores=1.2853739, pem=-4.5189333,
similarity=67.0225]
33: 2456377 National_League_Central
    [pem_pos=0, final_scores=0.56354934,
final_local_scores=0.58181673, pem=-0.98951703,
similarity=26.833433]
34: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.348279,
final_local_scores=1.3595862, pem=-4.650927,
similarity=69.95495]
35: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.0595621,
final_local_scores=1.0440106, pem=-3.771247,
similarity=55.753696]
36: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.3174921,
final_local_scores=1.3236489, pem=-4.03702,
similarity=65.54398]
37: 2456474 National_League_West
    [pem_pos=1, final_scores=0.5818624,
final_local_scores=0.5892046, pem=-1.2952092,
similarity=28.716303]
38: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.92519397,
final_local_scores=0.90168697, pem=-3.9307637,
similarity=52.373955]
39: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.1082902,
final_local_scores=1.1047587, pem=-4.482383,
similarity=61.43381]
40: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.96821636,
final_local_scores=0.9428199, pem=-4.409061,
similarity=56.2025]
41: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=1.092526,
final_local_scores=1.0783494, pem=-4.136474,
similarity=58.7646]
42: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.3164185,
final_local_scores=1.3247836, pem=-4.650927,
similarity=68.91637]
43: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.55621165,
final_local_scores=0.49455875, pem=-5.0692368,
similarity=46.41563]
44: 18110 Los_Angeles
    [pem_pos=0, final_scores=0.16510464,
```

```
final_local_scores=0.19479525, pem=-0.100652285,
similarity=10.450154]
45: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.9069986,
final_local_scores=0.87594885, pem=-4.409061,
similarity=54.206932]
46: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.0969322,
final_local_scores=1.0848317, pem=-3.771247,
similarity=56.97188]
47: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.2627484,
final_local_scores=1.2638497, pem=-4.03702,
similarity=63.759453]
48: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.0552365,
final_local_scores=1.0468056, pem=-4.482383,
similarity=59.704372]
49: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.2973037,
final_local_scores=1.3038571, pem=-4.5189333,
similarity=67.57407]
50: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.7572636,
final_local_scores=0.7182486, pem=-3.9307637,
similarity=46.899788]
1: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.23179518,
final_local_scores=0.14632392, pem=-4.422637,
similarity=32.507298]
2: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.34165913,
final_local_scores=0.26633358, pem=-4.422637,
similarity=36.088627]
3: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.14735784,
final_local_scores=0.04796976, pem=-5.0692368,
similarity=33.088524]
4: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.382079,
final_local_scores=0.3076812, pem=-4.5098605,
similarity=37.796856]
```

Sample: b'264_BASEBALL'

```
BASEBALL - [1 1 MAJOR LEAGUE 1] RESULTS FRIDAY . [2 2 NEW YO
RK 2] 1996-08-24 Results of [3 3 Major League 3] Baseball
games played on Friday ( home team in CAPS ) : [4 4 American L
eague 4] [5 5 BOSTON 5] 6 [6 6 Seattle 6] 4 [7 7 Milwaukee 7]
6 [8 8 CLEVELAND 8] 5 ( 11 innings ) 1 California 2 [9 9 BALT
IMORE 9] 0 [10 10 NEW YORK 10] 5 [11 11 Oakland 11] 3 [12 12 T
oronto 12] 4 2 CHICAGO 2 [13 13 Detroit 13] 3 [14 14 KANSAS CI
TY 14] 2 [15 15 MINNESOTA 15] 9 [16 16 Texas 16] 2 [17 17 Nat
```

ional League 17] [18 18 Cincinnati 18] 6 3 FLORIDA 5 (1ST GM) 4 FLORIDA 8 [19 19 Cincinnati 19] 3 (2ND GM) [20 20 AT LANTA 20] 4 5 Chicago 3 St Louis 1 [21 21 HOUSTON 21] 0 [22 22 Pittsburgh 22] 5 [23 23 COLORADO 23] 3 6 LOS ANGELES 7 [24 24 New York 24] 5 [25 25 Philadelphia 25] 7 [26 26 SAN DIEGO 26] 4 [27 27 Montreal 27] 10 [28 28 SAN FRANCISCO 28] 8

1: 38776 Major_League_Baseball
2: 645042 New_York_City
3: 38776 Major_League_Baseball
4: 64777 American_League
5: 4339 Boston_Red_Sox
6: 27660 Seattle_Mariners
7: 20152 Milwaukee_Brewers
8: 6652 Cleveland_Indians
9: 4340 Baltimore_Orioles
10: 4848143 New_York_Yankees
11: 22523 Oakland_Athletics
12: 30847 Toronto_Blue_Jays
13: 8579 Detroit_Tigers
14: 16883 Kansas_City_Royals
15: 20050 Minnesota_Twins
16: 30857 Texas_Rangers_(baseball)
17: 64290 National_League
18: 6671 Cincinnati_Reds
19: 6671 Cincinnati_Reds
20: 2140 Atlanta_Braves
21: 13894 Houston_Astros
22: 24270 Pittsburgh_Pirates
23: 6669 Colorado_Rockies
24: 4848143 New_York_Yankees
25: 23741 Philadelphia_Phillies
26: 28417 San_Diego_Padres
27: 20153 Montreal_Expos
28: 28416 San_Francisco_Giants

1: 1360083 Los_Angeles_Angels_of_Anheim
[winner: 7590031 California_Golden_Bears_football,
pem_pos=1, final_scores=-0.70514196,
final_local_scores=-0.77207893, pem=-4.4771376,
similarity=5.396708]
RECALL MISS: 1360083 Los_Angeles_Angels_of_Anheim]

2: 5945 Chicago_White_Sox
[winner: 994449 Chicago_(album), pem_pos=0,
final_scores=-0.70285815, final_local_scores=-0.63986725,
pem=0.0, similarity=-15.005201]
RECALL MISS: 5945 Chicago_White_Sox]

3: 11278 Miami_Marlins
[winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.78104997, final_local_scores=-0.85166055,
pem=-4.620281, similarity=3.8002708]
RECALL MISS: 11278 Miami_Marlins]

4: 11278 Miami_Marlins
[winner: 70551 Florida_Panthers, pem_pos=5,

```
final_scores=-0.72525096, final_local_scores=-0.7907086,
pem=-4.620281, similarity=5.619198|
    RECALL MISS: 11278 Miami_Marlins]
5: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.60611427, final_local_scores=-0.6821057,
pem=-4.591283, similarity=8.702428|
    RECALL MISS: 5945 Chicago_White_Sox]
6: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 18213 Los_Angeles_Dodgers, pem_pos=15,
final_scores=-0.23901762, final_local_scores=-0.37622494,
pem=-4.7524934, similarity=18.707209|
    RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
1: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.5507862,
final_local_scores=0.49693668, pem=-0.47792417,
similarity=21.51832]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.25401503,
final_local_scores=0.3191108, pem=-1.7906687,
similarity=23.35055]
3: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.56062835,
final_local_scores=0.5076877, pem=-0.47792417,
similarity=21.839151]
4: 64777 American_League
    [pem_pos=0, final_scores=0.5427007,
final_local_scores=0.510899, pem=0.0, similarity=19.335958]
5: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.0374286,
final_local_scores=1.0213768, pem=-4.418931,
similarity=58.600468]
6: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.4169827,
final_local_scores=0.35045558, pem=-3.8567436,
similarity=35.521576]
7: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.5458566,
final_local_scores=0.4847631, pem=-3.40205,
similarity=37.05689]
8: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.1415348,
final_local_scores=1.1293225, pem=-4.191953,
similarity=60.587444]
9: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.1766734,
final_local_scores=1.1683562, pem=-4.046561,
similarity=60.961617]
10: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.5435681,
final_local_scores=0.48092204, pem=-5.0692368,
similarity=46.008682]
```

```
11: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.136724,
final_local_scores=1.1259208, pem=-3.9188845,
similarity=59.00094]
12: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.83276373,
final_local_scores=0.80408883, pem=-4.422637,
similarity=52.136314]
13: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.0673999,
final_local_scores=1.0491855, pem=-4.1401463,
similarity=57.91426]
14: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.65526897,
final_local_scores=0.6045644, pem=-3.5915844,
similarity=41.662716]
15: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=1.1517439,
final_local_scores=1.143331, pem=-4.2906566,
similarity=61.542248]
16: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.0036992,
final_local_scores=0.98460203, pem=-4.928425,
similarity=60.273743]
17: 64290 National_League
    [pem_pos=0, final_scores=0.5918438,
final_local_scores=0.55911195, pem=-0.07431644,
similarity=21.178873]
18: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.1500446,
final_local_scores=1.1397893, pem=-3.771247,
similarity=58.611927]
19: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.2058029,
final_local_scores=1.200697, pem=-3.771247,
similarity=60.42953]
20: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=0.96651703,
final_local_scores=0.9395316, pem=-4.136474,
similarity=54.622]
21: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.1305809,
final_local_scores=1.1213744, pem=-4.650927,
similarity=62.84623]
22: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.0237262,
final_local_scores=1.00074, pem=-4.03702,
similarity=55.907734]
23: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.0268323,
final_local_scores=1.016073, pem=-4.482383,
similarity=58.78725]
```

```
24: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.53239465,
final_local_scores=0.4687168, pem=-5.0692368,
similarity=45.644455]
25: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.0432979,
final_local_scores=1.0240769, pem=-4.5189333,
similarity=59.224876]
26: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.85029775,
final_local_scores=0.8192605, pem=-3.9307637,
similarity=49.914185]
27: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.89306074,
final_local_scores=0.8668893, pem=-4.5098605,
similarity=54.484745]
28: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.8845144,
final_local_scores=0.8503149, pem=-4.409061,
similarity=53.441963]

Sample: b'265_SOCCER'
SOCCER - [1 1 PORTUGUESE 1] FIRST DIVISION RESULT . [2 2 LIS
BON 2] 1996-08-24 Result of a [3 3 Portuguese 3] first
division soccer match on Saturday : [4 4 Belenenses 4] 2 [5 5
Boavista 5] 4
1: 23033 Portugal
2: 18091 Lisbon
3: 23033 Portugal
4: 1629082 C.F._Os_Belenenses
5: 927845 Boavista_F.C.
1: 23033 Portugal
    [pem_pos=1, final_scores=0.29407305,
final_local_scores=0.24088418, pem=-1.127839,
similarity=17.411541]
2: 18091 Lisbon
    [pem_pos=0, final_scores=0.6504838,
final_local_scores=0.6508556, pem=-0.09749677,
similarity=24.042746]
3: 23033 Portugal
    [pem_pos=1, final_scores=0.62204033,
final_local_scores=0.5991387, pem=-1.127839,
similarity=28.10257]
4: 1629082 C.F._Os_Belenenses
    [pem_pos=0, final_scores=0.38532245,
final_local_scores=0.38524067, pem=0.0, similarity=15.586063]
5: 927845 Boavista_F.C.
    [pem_pos=0, final_scores=0.28184175,
final_local_scores=0.282751, pem=0.0, similarity=12.527567]
```

```
Sample: b'266_SOCCER'
SOCCER - DISAPPOINTING 1 AJAX SLUMP 2-0 AT [1 1 HEERENVEEN 1]
```

. [2 2 AMSTERDAM 2] 1996-08-24 [1 2 Dutch 1] champions [3 3 Ajax Amsterdam 3] faltered in their second league match of the season on Saturday losing 2-0 away at [4 4 Heerenveen 4] . [5 5 Ajax 5] , who had a dismal series of pre-season results before beating [6 6 NAC 6] of [7 7 Breda 7] in their opening game , had the best of an entertaining first half but failed to break the deadlock . Eight minutes after the interval , [8 8 Heerenveen 8] 's Romeo Wouden broke through the Amsterdam defence , left defender [9 9 John Veldman 9] standing and curled the ball beyond goalkeeper [10 10 Edwin van der Sar 10] into the [11 11 Ajax 11] net . [12 12 Ajax 12] , without injured defenders [13 13 Márcio Santos 13] and [14 14 Winston Bogarde 14] and strikers [15 15 Jari Litmanen 15] and [16 16 Marc Overmars 16] , then stepped up the pace and looked certain to equalise . But they left gaps at the back and on 73 minutes [17 17 Danish 17] striker [18 18 Jon Dahl Tomasson 18] rushed out of his own half , beat the [19 19 Ajax 19] defence and lobbed [20 20 van der Sar 20] . The defeat means [21 21 Ajax 21] 's main title contenders [22 22 PSV Eindhoven 22] , who beat the champions 3-0 in the traditional league curtain-raiser , can go three points clear of their rivals if they beat [23 23 Groningen 23] on Sunday .

1: 745757 SC_Heerenveen
2: 844 Amsterdam
3: 2273 AFC_Ajax
4: 745757 SC_Heerenveen
5: 2273 AFC_Ajax
6: 505173 NAC_Breda
7: 1098285 Breda
8: 745757 SC_Heerenveen
9: 7363720 John_Veldman
10: 685906 Edwin_van_der_Sar
11: 2273 AFC_Ajax
12: 2273 AFC_Ajax
13: 5960297 Márcio_Roberto_dos_Santos
14: 3109995 Winston_Bogarde
15: 539123 Jari_Litmanen
16: 645628 Marc_Overmars
17: 76972 Denmark
18: 735169 Jon_Dahl_Tomasson
19: 2273 AFC_Ajax
20: 685906 Edwin_van_der_Sar
21: 2273 AFC_Ajax
22: 79820 PSV_Eindhoven
23: 1103112 FC_Groningen
1: 2273 AFC_Ajax
[winner: 18539770 Jong_Ajax, pem_pos=14,
final_scores=-0.46359324, final_local_scores=-0.5580919,
pem=-4.0073333, similarity=9.227638]
RECALL MISS: 2273 AFC_Ajax
2: 1332105 Eredivisie
[winner: 9647657 Netherlands_national_football_team,

```
pem_pos=11, final_scores=0.053011283,
final_local_scores=-0.021554947, pem=-4.675365,
similarity=28.871832]
    RECALL MISS: 1332105 Eredivisie]
1: 745757 SC_Heerenveen
    [pem_pos=1, final_scores=0.5005501,
final_local_scores=0.4540462, pem=-0.6351416,
similarity=21.093355]
2: 844 Amsterdam
    [pem_pos=0, final_scores=0.46742296,
final_local_scores=0.51258415, pem=-0.0642803,
similarity=19.735813]
3: 2273 AFC_Ajax
    [pem_pos=0, final_scores=0.6709028,
final_local_scores=0.6203465, pem=0.0, similarity=22.602089]
4: 745757 SC_Heerenveen
    [pem_pos=1, final_scores=0.5685375,
final_local_scores=0.528312, pem=-0.6351416,
similarity=23.309597]
5: 2273 AFC_Ajax
    [pem_pos=1, final_scores=0.53126144,
final_local_scores=0.4678095, pem=-1.4356393,
similarity=25.857311]
6: 505173 NAC_Breda
    [pem_pos=1, final_scores=0.6359717,
final_local_scores=0.6178612, pem=-1.1774325,
similarity=28.930984]
7: 1098285 Breda
    [pem_pos=0, final_scores=0.48105162,
final_local_scores=0.55243903, pem=-0.08950387,
similarity=21.062332]
8: 745757 SC_Heerenveen
    [pem_pos=1, final_scores=0.5089596,
final_local_scores=0.46323222, pem=-0.6351416,
similarity=21.367485]
9: 7363720 John_Veldman
    [pem_pos=0, final_scores=0.69464177,
final_local_scores=0.79038566, pem=0.0, similarity=27.676397]
10: 685906 Edwin_van_der_Sar
    [pem_pos=0, final_scores=0.7138983,
final_local_scores=0.7446964, pem=0.0, similarity=26.312939]
11: 2273 AFC_Ajax
    [pem_pos=1, final_scores=0.61570364,
final_local_scores=0.56004983, pem=-1.4356393,
similarity=28.609945]
12: 2273 AFC_Ajax
    [pem_pos=1, final_scores=0.5641082,
final_local_scores=0.5036896, pem=-1.4356393,
similarity=26.928045]
13: 5960297 Márcio_Roberto_dos_Santos
    [pem_pos=0, final_scores=0.4161479,
final_local_scores=0.51727074, pem=0.0, similarity=19.526104]
```

```
14: 3109995 Winston_Bogarde
    [pem_pos=0, final_scores=0.7164399,
final_local_scores=0.7577773, pem=0.0, similarity=26.703299]
15: 539123 Jari_Litmanen
    [pem_pos=0, final_scores=0.8777142,
final_local_scores=0.92653674, pem=0.0, similarity=31.739418]
16: 645628 Marc_Overmars
    [pem_pos=0, final_scores=0.80580264,
final_local_scores=0.84384197, pem=0.0, similarity=29.27164]
17: 76972 Denmark
    [pem_pos=0, final_scores=0.5193547,
final_local_scores=0.61826986, pem=-0.65021485,
similarity=26.076088]
18: 735169 Jon_Dahl_Tomasson
    [pem_pos=0, final_scores=0.6477687,
final_local_scores=0.6738187, pem=0.0, similarity=24.197807]
19: 2273 AFC_Ajax
    [pem_pos=1, final_scores=0.6504727,
final_local_scores=0.5980297, pem=-1.4356393,
similarity=29.74334]
20: 685906 Edwin_van_der_Sar
    [pem_pos=0, final_scores=0.53657657,
final_local_scores=0.55099934, pem=0.0, similarity=20.532633]
21: 2273 AFC_Ajax
    [pem_pos=1, final_scores=0.6652364,
final_local_scores=0.6141568, pem=-1.4356393,
similarity=30.224604]
22: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.569036,
final_local_scores=0.51975393, pem=0.0, similarity=19.600206]
23: 1103112 FC_Groningen
    [pem_pos=2, final_scores=0.57942754,
final_local_scores=0.5470565, pem=-1.9677447,
similarity=31.115871]
1: 9647657 Netherlands_national_football_team
    [pem_pos=11, final_scores=0.053011283,
final_local_scores=-0.021554947, pem=-4.675365,
similarity=28.871832]
```

Sample: b'267_SOCCER'

```
SOCCER - [1 1 BELGIAN 1] FIRST DIVISION RESULTS . [2 2 BRUSS
ELS 2] 1996-08-24 Results of [3 3 Belgian 3] first division
matches on Saturday : [4 4 Standard Liege 4] 3 [5 5 Molenbeek
5] 0 [6 6 Anderlecht 6] 2 [7 7 Lokeren 7] 2 [8 8 Cercle Brugge
8] 2 [9 9 Mouscron 9] 2 1 Antwerp 1 [10 10 Lommel 10] 4 [11 11
Ghent 11] 3 [12 12 Aalst 12] 2 [13 13 Lierse 13] 4 [14 14 Cha
rleroi 14] 0 [15 15 Sint Truiden 15] 3 Ekeren 1
1: 3343 Belgium
2: 3708 Brussels
3: 3343 Belgium
4: 990794 Standard_Liège
5: 1452159 R._White_Daring_Molenbeek
```

6: 26584259 R.S.C._Anderlecht
7: 1437580 K.S.C._Lokeren_Oost-Vlaanderen
8: 1434036 Cercle_Brugge_K.S.V.
9: 1434546 R.E._Mouscron
10: 3470620 K.F.C._Lommel_S.K.
11: 1521107 K.A.A._Gent
12: 1537131 SC_Eendracht_Aalst
13: 842033 Lierse_S.K.
14: 1433845 R.Charleroi_S.C.
15: 1437246 K._Sint-Truidense_V.V.
1: 1520840 Royal_Antwerp_FC
[winner: 2107074 Royal_Antwerp_F.C., pem_pos=2,
final_scores=-0.04852447, final_local_scores=-0.13969725,
pem=-2.9206293, similarity=15.803703]
RECALL MISS: 1520840 Royal_Antwerp_FC]
1: 3343 Belgium
[pem_pos=0, final_scores=0.5806283,
final_local_scores=0.61593926, pem=-0.12403572,
similarity=23.145094]
2: 3708 Brussels
[pem_pos=0, final_scores=0.5328661,
final_local_scores=0.589569, pem=-0.025440974,
similarity=21.82198]
3: 3343 Belgium
[pem_pos=0, final_scores=0.61208326,
final_local_scores=0.6502991, pem=-0.12403572,
similarity=24.17046]
4: 990794 Standard_Liège
[pem_pos=0, final_scores=0.51088923,
final_local_scores=0.4692346, pem=0.0, similarity=18.09261]
5: 1452159 R.White_Daring_Molenbeek
[pem_pos=0, final_scores=0.41630584,
final_local_scores=0.3929445, pem=-0.005970144,
similarity=15.848428]
6: 26584259 R.S.C._Anderlecht
[pem_pos=0, final_scores=0.630871,
final_local_scores=0.5955051, pem=0.0, similarity=21.860773]
7: 1437580 K.S.C._Lokeren_Oost-Vlaanderen
[pem_pos=0, final_scores=0.54556924,
final_local_scores=0.51769227, pem=0.0, similarity=19.538683]
8: 1434036 Cercle_Brugge_K.S.V.
[pem_pos=0, final_scores=0.5106724,
final_local_scores=0.46884918, pem=0.0, similarity=18.081108]
9: 1434546 R.E._Mouscron
[pem_pos=1, final_scores=0.48739558,
final_local_scores=0.4797861, pem=-0.28243667,
similarity=19.943422]
10: 3470620 K.F.C._Lommel_S.K.
[pem_pos=1, final_scores=0.42259508,
final_local_scores=0.44565874, pem=-1.6453731,
similarity=26.336853]
11: 1521107 K.A.A._Gent

```
[pem_pos=13, final_scores=0.5503583,
final_local_scores=0.5193266, pem=-2.9399738,
similarity=35.57549]
12: 1537131 SC_Eendracht_Aalst
    [pem_pos=1, final_scores=0.4888224,
final_local_scores=0.4708789, pem=-1.4374129,
similarity=25.958553]
13: 842033 Lierse_S.K.
    [pem_pos=0, final_scores=0.54108757,
final_local_scores=0.5063132, pem=0.0, similarity=19.199108]
14: 1433845 R._Charleroi_S.C.
    [pem_pos=1, final_scores=0.5815093,
final_local_scores=0.5561904, pem=-1.0632312,
similarity=26.46956]
15: 1437246 K._Sint-Truidense_V.V.
    [pem_pos=0, final_scores=0.5698759,
final_local_scores=0.55118257, pem=0.0, similarity=20.5381]
```

Sample: b'268_SOCCER'

```
SOCCER - LEADING [1 1 FRENCH 1] LEAGUE SCORERS . [2 2 PARIS 2
] 1996-08-24 Leading goalscorers in the [3 3 French 3] first
division after Saturday 's matches : 3 - [4 4 Anto Drobnjak 4]
( [5 5 Bastia 5] ) , [6 6 Xavier Gravelaine 6] ( [7 7 Marseil
le 7] ) . 2 - [8 8 Miladin Becanovic 8] ( [9 9 Lille 9] ) , [1
0 10 Enzo Scifo 10] ( [11 11 Monaco 11] ) , [12 12 Vladimír Sm
icer 12] ( [13 13 Lens 13] ) , [14 14 Christopher Wreh 14] ( [
15 15 Guingamp 15] ) .
```

```
1: 5843419 France
2: 22989 Paris
3: 5843419 France
4: 8809288 Anto_Drobnjak
5: 1256442 SC_Bastia
6: 10326208 Xavier_Gravelaine
7: 492068 Olympique_de_Marseille
8: 13449225 Miladin_Bečanović
9: 990757 Lille(OSC)
10: 52265 Enzo_Scifo
11: 222981 AS_Monaco_FC
12: 549482 Vladimír_Smicer
13: 1075697 RC_Lens
14: 2839275 Christopher_Wreh
15: 1693070 En_Avant_de_Guingamp
1: 5843419 France
    [pem_pos=0, final_scores=0.6452647,
final_local_scores=0.7002182, pem=-0.80223274,
similarity=29.348288]
2: 22989 Paris
    [pem_pos=0, final_scores=0.51445013,
final_local_scores=0.56952506, pem=-0.115301244,
similarity=21.712503]
3: 5843419 France
    [pem_pos=0, final_scores=0.594778,
```

```
final_local_scores=0.6450692, pem=-0.80223274,
similarity=27.70253]
4: 8809288 Anto_Drobnjak
    [pem_pos=0, final_scores=0.88318974,
final_local_scores=1.0059919, pem=0.0, similarity=34.110516]
5: 1256442 SC_Bastia
    [pem_pos=1, final_scores=0.4497732,
final_local_scores=0.4144178, pem=-0.7022244,
similarity=20.275572]
6: 10326208 Xavier_Gravelaine
    [pem_pos=0, final_scores=0.6486216,
final_local_scores=0.7398362, pem=0.0, similarity=26.167902]
7: 492068 Olympique_de_Marseille
    [pem_pos=1, final_scores=0.64819884,
final_local_scores=0.6136911, pem=-1.4345005,
similarity=30.204517]
8: 13449225 Miladin_Bečanović
    [pem_pos=0, final_scores=0.3624639,
final_local_scores=0.5557176, pem=0.0, similarity=20.673433]
9: 990757 Lille OSC
    [pem_pos=1, final_scores=0.6092586,
final_local_scores=0.58077997, pem=-1.3501564,
similarity=28.763706]
10: 52265 Enzo_Scifo
    [pem_pos=0, final_scores=0.97692984,
final_local_scores=1.0572581, pem=0.0, similarity=35.640408]
11: 222981 AS_Monaco_FC
    [pem_pos=1, final_scores=0.6261907,
final_local_scores=0.60167366, pem=-2.451784,
similarity=35.378036]
12: 549482 Vladimír_Šmicer
    [pem_pos=0, final_scores=0.8383785,
final_local_scores=0.9469773, pem=0.0, similarity=32.349403]
13: 1075697 RC_Lens
    [pem_pos=0, final_scores=0.4485125,
final_local_scores=0.40631908, pem=-0.86081105,
similarity=20.896309]
14: 2839275 Christopher_Wreh
    [pem_pos=0, final_scores=0.60534734,
final_local_scores=0.6824878, pem=0.0, similarity=24.45651]
15: 1693070 En_Avant_de_Guingamp
    [pem_pos=0, final_scores=0.36820227,
final_local_scores=0.34141654, pem=0.0, similarity=14.278263]
```

Sample: b'269_SOCCER'
SOCCER - [1 1 FRENCH 1] LEAGUE SUMMARIES . [2 2 PARIS 2]
1996-08-24 Summaries of [3 3 French 3] first division matches
on Saturday : [4 4 Nantes 4] 0 [5 5 Lens 5] 1 ([6 6 Smicer 6]
52nd) . Halftime 0-0 . Attendance 16,000 . [7 7 Nice 7] 1 ([
8 8 Debbah 8] 39th) [9 9 Bastia 9] 1 ([10 10 Drobnjak 10]
82nd) . 1-0 . 6,000 . [11 11 Lille 11] 3 (Boutoille 47th ,
Bejanovic 79th pen , 82nd)) [12 12 Rennes 12] 1 (Guivarc'h

60th pen .) 0-0 . 6,000 . [13 13 Bordeaux 13] 0 [14 14 Auxerre 14] 0 . 30,000 . [15 15 Marseille 15] 1 ([16 16 Gravelaine 16] 24th) [17 17 Metz 17] 2 (Traore 65th , [18 18 Bombarda 18] 69th) . 1-0 . 20,000 . [19 19 Strasbourg 19] 1 ([20 20 Zitelli 20] 80th) [21 21 Le Havre 21] 0 . 0-0 . 15,000 [22 22 Caen 22] 1 ([23 23 Bancarel 23] 70th) [24 24 Lyon 24] 1 ([25 25 Caveglia 25] 89th) . 0-0 . 9,000 . [26 26 Guingamp 26] 2 ([27 27 Wreh 27] 15th , 42nd) [28 28 Monaco 28] 1 ([29 29 Scifo 29] 35th) . 2-1 . 7,000 . [30 30 Montpellier 30] 0 [31 31 Cannes 31] 1 ([32 32 Charvet 32] 8th) . 0-1 . 10,000 .
Played on Friday : [33 33 Nancy 33] 0 [34 34 Paris St Germain 34] 0 . 15,000 .

1: 5843419 France
2: 22989 Paris
3: 5843419 France
4: 1135276 FC_Nantes
5: 1075697 RC_Lens
6: 549482 Vladimír_Šmicer
7: 704478 OGC_Nice
8: 11527631 James_Debbah
9: 1256442 SC_Bastia
10: 8809288 Anto_Drobnjak
11: 990757 Lille OSC
12: 1004187 Stade_Rennais_F.C.
13: 1135285 FC_Girondins_de_Bordeaux
14: 990753 AJ_Auxerre
15: 492068 Olympique_de_Marseille
16: 10326208 Xavier_Gravelaine
17: 1135310 FC_Metz
18: 9425024 Mariano_Bombarda
19: 1256446 RC_Strasbourg
20: 15098218 David_Zitelli
21: 3658991 Le_Havre_AC
22: 1256445 Stade_Malherbe_Caen
23: 15273188 Anthony_Bancarel
24: 979521 Olympique_Lyonnais
25: 11576169 Alain_Caveglia
26: 1693070 En_Avant_de_Guingamp
27: 2839275 Christopher_Wreh
28: 222981 AS_Monaco_FC
29: 52265 Enzo_Scifo
30: 1640865 Montpellier_HSC
31: 2237150 AS_Cannes
32: 7044971 Laurent_Charvet
33: 1640855 AS_Nancy
34: 357488 Paris_Saint-Germain_F.C.

1: 5843419 France
[pem_pos=0, final_scores=0.65639466,
final_local_scores=0.7296904, pem=-0.80223274,
similarity=30.227797]
2: 22989 Paris
[pem_pos=0, final_scores=0.5191724,

```
final_local_scores=0.58545005, pem=-0.115301244,
similarity=22.187737]
3: 5843419 France
    [pem_pos=0, final_scores=0.68612856,
final_local_scores=0.76217014, pem=-0.80223274,
similarity=31.197056]
4: 1135276 FC_Nantes
    [pem_pos=1, final_scores=0.6428866,
final_local_scores=0.60550207, pem=-1.4553026,
similarity=30.073265]
5: 1075697 RC_Lens
    [pem_pos=0, final_scores=0.5068545,
final_local_scores=0.45562768, pem=-0.86081105,
similarity=22.367777]
6: 549482 Vladimír Šmicer
    [pem_pos=0, final_scores=0.5019085,
final_local_scores=0.5733648, pem=0.0, similarity=21.200062]
7: 704478 OGC_Nice
    [pem_pos=1, final_scores=0.89482623,
final_local_scores=0.88528115, pem=-2.1390686,
similarity=42.140854]
8: 11527631 James_Debbah
    [pem_pos=0, final_scores=0.50343734,
final_local_scores=0.57315046, pem=0.0, similarity=21.193666]
9: 1256442 SC_Bastia
    [pem_pos=1, final_scores=0.49686676,
final_local_scores=0.45240915, pem=-0.7022244,
similarity=21.40931]
10: 8809288 Anto_Drobnjak
    [pem_pos=2, final_scores=0.48329765,
final_local_scores=0.57004607, pem=-0.55814195,
similarity=24.136286]
11: 990757 Lille OSC
    [pem_pos=1, final_scores=0.6377161,
final_local_scores=0.5983073, pem=-1.3501564,
similarity=29.286758]
12: 1004187 Stade_Rennais_F.C.
    [pem_pos=1, final_scores=0.4122942,
final_local_scores=0.35265404, pem=-1.0138907,
similarity=20.127308]
13: 1135285 FC_Girondins_de_Bordeaux
    [pem_pos=1, final_scores=0.67147493,
final_local_scores=0.6311987, pem=-1.6081072,
similarity=31.671078]
14: 990753 AJ_Auxerre
    [pem_pos=0, final_scores=0.63031596,
final_local_scores=0.59055597, pem=-0.30518255,
similarity=23.372711]
15: 492068 Olympique_de_Marseille
    [pem_pos=1, final_scores=0.6406825,
final_local_scores=0.59481466, pem=-1.4345005,
similarity=29.641205]
```

```
16: 10326208 Xavier_Gravelaine
    [pem_pos=0, final_scores=0.47937,
final_local_scores=0.54864776, pem=0.0, similarity=20.462456]
17: 1135310 FC_Metz
    [pem_pos=1, final_scores=0.7500377,
final_local_scores=0.72550946, pem=-1.9220247,
similarity=36.19263]
18: 9425024 Mariano_Bombarda
    [pem_pos=2, final_scores=0.336376,
final_local_scores=0.44661558, pem=-0.6931472,
similarity=21.187054]
19: 1256446 RC_Strasbourg
    [pem_pos=1, final_scores=0.67151153,
final_local_scores=0.6444426, pem=-2.2163424,
similarity=35.37398]
20: 15098218 David_Zitelli
    [pem_pos=0, final_scores=0.575525,
final_local_scores=0.701822, pem=0.0, similarity=25.033482]
21: 3658991 Le_Havre_AC
    [pem_pos=1, final_scores=0.4579693,
final_local_scores=0.4136654, pem=-1.3456037,
similarity=23.751915]
22: 1256445 Stade_Malherbe_Caen
    [pem_pos=1, final_scores=0.39838594,
final_local_scores=0.34800166, pem=-1.6168041,
similarity=23.267208]
23: 15273188 Anthony_Bancarel
    [pem_pos=0, final_scores=0.36629105,
final_local_scores=0.46543908, pem=0.0, similarity=17.979343]
24: 979521 Olympique_Lyonnais
    [pem_pos=1, final_scores=0.7702146,
final_local_scores=0.7417595, pem=-1.9193003,
similarity=36.662746]
25: 11576169 Alain_Caveglia
    [pem_pos=0, final_scores=0.582995,
final_local_scores=0.6589397, pem=0.0, similarity=23.753788]
26: 1693070 En_Avant_de_Guingamp
    [pem_pos=0, final_scores=0.4457113,
final_local_scores=0.4096318, pem=0.0, similarity=16.313942]
27: 2839275 Christopher_Wreh
    [pem_pos=0, final_scores=0.51509005,
final_local_scores=0.5703492, pem=0.0, similarity=21.110071]
28: 222981 AS_Monaco_FC
    [pem_pos=1, final_scores=0.8442746,
final_local_scores=0.82569844, pem=-2.451784,
similarity=42.063385]
29: 52265 Enzo_Scifo
    [pem_pos=0, final_scores=0.60676533,
final_local_scores=0.6444094, pem=0.0, similarity=23.320175]
30: 1640865 Montpellier_HSC
    [pem_pos=1, final_scores=0.5822504,
final_local_scores=0.5436114, pem=-1.476352,
```

```

similarity=28.340794]
31: 2237150 AS_Cannes
    [pem_pos=2, final_scores=0.77806216,
final_local_scores=0.7715067, pem=-2.2011297,
similarity=39.083096]
32: 7044971 Laurent_Charvet
    [pem_pos=5, final_scores=0.40872318,
final_local_scores=0.4712687, pem=-1.7917595,
similarity=27.897179]
33: 1640855 AS_Nancy
    [pem_pos=0, final_scores=0.8514057,
final_local_scores=0.8348412, pem=-1.6882259,
similarity=38.183876]
34: 357488 Paris_Saint-Germain_F.C.
    [pem_pos=0, final_scores=0.5147703,
final_local_scores=0.46178377, pem=0.0, similarity=17.870262]

```

Sample: b'270_SOCCER'

SOCCER - [1 1 FRENCH 1] LEAGUE STANDINGS . [2 2 PARIS 2]
1996-08-24 Standings in the [3 3 French 3] soccer league
after Saturday 's matches (tabulate under played , won ,
drawn , lost , goals for , goals against , points) : [4 4 Le
ns 4] 3 3 0 0 6 1 9 [5 5 Bastia 5] 3 2 1 0 4 1 7 [6 6 Paris Sa
int-Germain 6] 3 2 1 0 3 0 7 [7 7 Auxerre 7] 3 2 1 0 3 0 7 [8
8 Cannes 8] 3 2 1 0 4 2 7 [9 9 Lille 9] 3 2 0 1 4 3 6 [10 10 B
ordeaux 10] 3 1 2 0 2 1 5 [11 11 Monaco 11] 3 1 1 1 5 3 4 [12
12 Marseille 12] 3 1 1 1 5 4 4 [13 13 Metz 13] 3 1 1 1 3 3 4 [
14 14 Lyon 14] 3 1 1 1 4 4 4 [15 15 Guingamp 15] 3 1 1 1 2 2
4 [16 16 Rennes 16] 3 1 0 2 4 6 3 [17 17 Strasbourg 17] 3 1 0
2 1 3 3 [18 18 Montpellier 18] 3 0 2 1 1 2 2 [19 19 Nantes 19]
3 0 1 2 2 5 1 [20 20 Nancy 20] 3 0 1 2 2 5 1 [21 21 Nice 21]
3 0 1 2 2 5 1 [22 22 Le Havre 22] 3 0 1 1 1 3 1 [23 23 Caen 23
] 3 0 1 2 1 5 1

- 1: 5843419 France
- 2: 22989 Paris
- 3: 5843419 France
- 4: 1075697 RC_Lens
- 5: 1256442 SC_Bastia
- 6: 357488 Paris_Saint-Germain_F.C.
- 7: 990753 AJ_Auxerre
- 8: 2237150 AS_Cannes
- 9: 990757 Lille OSC
- 10: 1135285 FC_Girondins_de_Bordeaux
- 11: 222981 AS_Monaco_FC
- 12: 492068 Olympique_de_Marseille
- 13: 1135310 FC_Metz
- 14: 979521 Olympique_Lyonnais
- 15: 1693070 En_Avant_de_Guingamp
- 16: 1004187 Stade_Rennais_F.C.
- 17: 1256446 RC_Strasbourg
- 18: 1640865 Montpellier_HSC
- 19: 1135276 FC_Nantes

20: 1640855 AS_Nancy
21: 704478 OGC_Nice
22: 3658991 Le_Havre_AC
23: 1256445 Stade_Malherbe_Caen
1: 5843419 France
[pem_pos=0, final_scores=0.64638925,
final_local_scores=0.71618754, pem=-0.80223274,
similarity=29.824844]
2: 22989 Paris
[pem_pos=0, final_scores=0.51925975,
final_local_scores=0.58089507, pem=-0.115301244,
similarity=22.051805]
3: 5843419 France
[pem_pos=0, final_scores=0.62404656,
final_local_scores=0.6917816, pem=-0.80223274,
similarity=29.096521]
4: 1075697 RC_Lens
[pem_pos=0, final_scores=0.5049754,
final_local_scores=0.4521957, pem=-0.86081105,
similarity=22.26536]
5: 1256442 SC_Bastia
[pem_pos=1, final_scores=0.43687326,
final_local_scores=0.3862155, pem=-0.7022244,
similarity=19.433958]
6: 357488 Paris_Saint-Germain_F.C.
[pem_pos=0, final_scores=0.48956174,
final_local_scores=0.43196762, pem=0.0, similarity=16.980488]
7: 990753 AJ_Auxerre
[pem_pos=0, final_scores=0.6216396,
final_local_scores=0.58045566, pem=-0.30518255,
similarity=23.071299]
8: 2237150 AS_Cannes
[pem_pos=2, final_scores=0.5964494,
final_local_scores=0.57007545, pem=-2.2011297,
similarity=33.071987]
9: 990757 Lille OSC
[pem_pos=1, final_scores=0.5620089,
final_local_scores=0.5127539, pem=-1.3501564,
similarity=26.733671]
10: 1135285 FC_Girondins_de_Bordeaux
[pem_pos=1, final_scores=0.5370647,
final_local_scores=0.48182088, pem=-1.6081072,
similarity=27.213345]
11: 222981 AS_Monaco_FC
[pem_pos=1, final_scores=0.7376222,
final_local_scores=0.70739, pem=-2.451784,
similarity=38.532825]
12: 492068 Olympique_de_Marseille
[pem_pos=1, final_scores=0.5942225,
final_local_scores=0.5409547, pem=-1.4345005,
similarity=28.033916]
13: 1135310 FC_Metz

```

[pem_pos=1, final_scores=0.6763992,
final_local_scores=0.64277095, pem=-1.9220247,
similarity=33.72355]
14: 979521 Olympique_Lyonnais
[pem_pos=1, final_scores=0.6580781,
final_local_scores=0.6184296, pem=-1.9193003,
similarity=32.982338]
15: 1693070 En_Avant_de_Guingamp
[pem_pos=0, final_scores=0.46643335,
final_local_scores=0.42831516, pem=0.0, similarity=16.871492]
16: 1004187 Stade_Rennais_F.C.
[pem_pos=1, final_scores=0.36221653,
final_local_scores=0.2965181, pem=-1.0138907,
similarity=18.452099]
17: 1256446 RC_Strasbourg
[pem_pos=1, final_scores=0.64350027,
final_local_scores=0.6112065, pem=-2.2163424,
similarity=34.38215]
18: 1640865 Montpellier_HSC
[pem_pos=1, final_scores=0.5380878,
final_local_scores=0.4916494, pem=-1.476352,
similarity=26.790142]
19: 1135276 FC_Nantes
[pem_pos=1, final_scores=0.56735677,
final_local_scores=0.5223069, pem=-1.4553026,
similarity=27.590555]
20: 1640855 AS_Nancy
[pem_pos=0, final_scores=0.70379275,
final_local_scores=0.6742242, pem=-1.6882259,
similarity=33.390743]
21: 704478 OGC_Nice
[pem_pos=1, final_scores=0.7623709,
final_local_scores=0.73603487, pem=-2.1390686,
similarity=37.68705]
22: 3658991 Le_Havre_AC
[pem_pos=1, final_scores=0.46661967,
final_local_scores=0.4192956, pem=-1.3456037,
similarity=23.919931]
23: 1256445 Stade_Malherbe_Caen
[pem_pos=1, final_scores=0.43143433,
final_local_scores=0.38156927, pem=-1.6168041,
similarity=24.268932]

```

Sample: b'271_SOCCER'

SOCCER - [1 1 FRENCH 1] LEAGUE RESULTS . [2 2 PARIS 2]

1996-08-24 Results of [3 3 French 3] first division matches

on Saturday : [4 4 Nantes 4] 0 [5 5 Lens 5] 1 [6 6 Nice 6] 1 [7 7 Bastia 7] 1 [8 8 Lille 8] 3 [9 9 Rennes 9] 1 [10 10 Bordeaux 10] 0 [11 11 Auxerre 11] 0 [12 12 Marseille 12] 1 [13 13 Metz 13] 2 [14 14 Strasbourg 14] 1 [15 15 Le Havre 15] 0 [16 16 Caen 16] 1 [17 17 Lyon 17] 1 [18 18 Guingamp 18] 2 [19 19 Monaco 19] 1 [20 20 Montpellier 20] 0 [21 21 Cannes 21] 1 Played

```
on Friday : [22 22 Nancy 22] 0 [23 23 Paris St Germain 23] 0
1: 5843419 France
2: 22989 Paris
3: 5843419 France
4: 1135276 FC_Nantes
5: 1075697 RC_Lens
6: 704478 OGC_Nice
7: 1256442 SC_Bastia
8: 990757 Lille_OSC
9: 1004187 Stade_Rennais_F.C.
10: 1135285 FC_Girondins_de_Bordeaux
11: 990753 AJ_Auxerre
12: 492068 Olympique_de_Marseille
13: 1135310 FC_Metz
14: 1256446 RC_Strasbourg
15: 3658991 Le_Havre_AC
16: 1256445 Stade_Malherbe_Caen
17: 979521 Olympique_Lyonnais
18: 1693070 En_Avant_de_Guingamp
19: 222981 AS_Monaco_FC
20: 1640865 Montpellier_HSC
21: 2237150 AS_Cannes
22: 1640855 AS_Nancy
23: 357488 Paris_Saint-Germain_F.C.
1: 5843419 France
    [pem_pos=0, final_scores=0.6679278,
final_local_scores=0.73971516, pem=-0.80223274,
similarity=30.526955]
2: 22989 Paris
    [pem_pos=0, final_scores=0.51275426,
final_local_scores=0.5737888, pem=-0.115301244,
similarity=21.839743]
3: 5843419 France
    [pem_pos=0, final_scores=0.7016449,
final_local_scores=0.77654594, pem=-0.80223274,
similarity=31.626059]
4: 1135276 FC_Nantes
    [pem_pos=1, final_scores=0.6637614,
final_local_scores=0.6276143, pem=-1.4553026,
similarity=30.733137]
5: 1075697 RC_Lens
    [pem_pos=0, final_scores=0.5282231,
final_local_scores=0.47759026, pem=-0.86081105,
similarity=23.023184]
6: 704478 OGC_Nice
    [pem_pos=1, final_scores=0.73826545,
final_local_scores=0.70970345, pem=-2.1390686,
similarity=36.901268]
7: 1256442 SC_Bastia
    [pem_pos=1, final_scores=0.46803993,
final_local_scores=0.42026043, pem=-0.7022244,
similarity=20.449926]
```

```
8: 990757 Lille OSC
    [pem_pos=1, final_scores=0.621844,
final_local_scores=0.57811457, pem=-1.3501564,
similarity=28.684164]
9: 1004187 Stade Rennais F.C.
    [pem_pos=1, final_scores=0.40125257,
final_local_scores=0.339159, pem=-1.0138907,
similarity=19.72459]
10: 1135285 FC Girondins de Bordeaux
    [pem_pos=1, final_scores=0.62992686,
final_local_scores=0.5832588, pem=-1.6081072,
similarity=30.240456]
11: 990753 AJ Auxerre
    [pem_pos=0, final_scores=0.63967496,
final_local_scores=0.60015655, pem=-0.30518255,
similarity=23.659212]
12: 492068 Olympique de Marseille
    [pem_pos=1, final_scores=0.5809613,
final_local_scores=0.5264689, pem=-1.4345005,
similarity=27.601631]
13: 1135310 FC Metz
    [pem_pos=1, final_scores=0.5975138,
final_local_scores=0.55660075, pem=-1.9220247,
similarity=31.15206]
14: 1256446 RC Strasbourg
    [pem_pos=1, final_scores=0.6075751,
final_local_scores=0.5719637, pem=-2.2163424,
similarity=33.211067]
15: 3658991 Le Havre AC
    [pem_pos=1, final_scores=0.43811315,
final_local_scores=0.38815653, pem=-1.3456037,
similarity=22.990679]
16: 1256445 Stade Malherbe Caen
    [pem_pos=1, final_scores=0.42156804,
final_local_scores=0.3707918, pem=-1.6168041,
similarity=23.947311]
17: 979521 Olympique Lyonnais
    [pem_pos=1, final_scores=0.7153872,
final_local_scores=0.68103117, pem=-1.9193003,
similarity=34.850494]
18: 1693070 En Avant de Guingamp
    [pem_pos=0, final_scores=0.49507612,
final_local_scores=0.459603, pem=0.0, similarity=17.805183]
19: 222981 AS Monaco FC
    [pem_pos=1, final_scores=0.7049296,
final_local_scores=0.67167836, pem=-2.451784,
similarity=37.467117]
20: 1640865 Montpellier HSC
    [pem_pos=1, final_scores=0.58848375,
final_local_scores=0.5466993, pem=-1.476352,
similarity=28.432941]
21: 2237150 AS Cannes
```

```

    [pem_pos=2, final_scores=0.74701244,
final_local_scores=0.7345428, pem=-2.2011297,
similarity=37.98002]
22: 1640855 AS_Nancy
    [pem_pos=0, final_scores=0.83106416,
final_local_scores=0.8132488, pem=-1.6882259,
similarity=37.539516]
23: 357488 Paris_Saint-Germain_F.C.
    [pem_pos=0, final_scores=0.5302489,
final_local_scores=0.47641212, pem=0.0, similarity=18.3068]

```

Sample: b'272_SOCCER'

SOCCER - [1 1 DUTCH 1] FIRST DIVISION RESULTS / STANDINGS . [2 2 AMSTERDAM 2] 1996-08-24 Result of [3 3 Dutch 3] first division soccer match played on Saturday : Graafschap Doetinchem 3 [4 4 RKC Waalwijk 4] 2 [5 5 Willem II Tilburg 5] 0 [6 6 Fortuna Sittard 6] 1 [7 7 NAC Breda 7] 1 [8 8 Sparta Rotterdam 8] 0 [9 9 Heerenveen 9] 2 [10 10 Ajax Amsterdam 10] 0 Standings (tabulated under played , won , drawn , lost , goals for , against , points) : Graafschap Doetinchem 2 1 1 0 4 3 4 [11 11 Fortuna Sittard 11] 2 1 1 0 1 0 4 [12 12 PSV Eindhoven 12] 1 1 0 0 4 1 3 [13 13 Twente Enschede 13] 1 1 0 0 3 1 3 [14 14 Vitesse Arnhem 14] 1 1 0 0 2 0 3 [15 15 Heerenveen 15] 2 1 0 1 3 3 3 [16 16 NAC Breda 16] 2 1 0 1 1 1 3 [17 17 Ajax Amsterdam 17] 2 1 0 1 1 2 3 [18 18 Utrecht 18] 1 0 1 0 2 2 1 [19 19 Feyenoord Rotterdam 19] 1 0 1 0 1 1 1 [20 20 Roda JC Kerkrade 20] 1 0 1 0 1 1 1 [21 21 Volendam 21] 1 0 1 0 1 1 1 1 [22 22 Groningen 22] 1 0 1 0 0 0 1 [23 23 RKC Waalwijk 23] 2 0 1 1 4 5 1 [24 24 Sparta Rotterdam 24] 2 0 1 1 0 1 1 [25 25 Willem II Tilburg 25] 2 0 1 1 0 1 1 [26 26 AZ Alkmaar 26] 1 0 0 1 0 2 0 [27 27 NEC Nijmegen 27] 1 0 0 1 1 4 0

1: 21148 Netherlands
2: 844 Amsterdam
3: 21148 Netherlands
4: 1107824 RKC_Waalwijk
5: 834275 Willem_II_(football_club)
6: 1336325 Fortuna_Sittard
7: 505173 NAC_Breda
8: 79818 Sparta_Rotterdam
9: 745757 SC_Heerenveen
10: 2273 AFC_Ajax
11: 1336325 Fortuna_Sittard
12: 79820 PSV_Eindhoven
13: 800033 FC_Twente
14: 834256 Vitesse
15: 745757 SC_Heerenveen
16: 505173 NAC_Breda
17: 2273 AFC_Ajax
18: 834226 FC_Utrecht
19: 47707 Feyenoord
20: 834196 Roda_JC_Kerkrade
21: 2202202 FC_Volendam

```
22: 1103112 FC_Groningen
23: 1107824 RKC_Waalwijk
24: 79818 Sparta_Rotterdam
25: 834275 Willem_II_(football_club)
26: 2289 AZ_Alkmaar
27: 834154 N.E.C._(football_club)
1: 21148 Netherlands
    [pem_pos=0, final_scores=0.6498272,
final_local_scores=0.711097, pem=-0.7943452,
similarity=29.63004]
2: 844 Amsterdam
    [pem_pos=0, final_scores=0.57904476,
final_local_scores=0.6421291, pem=-0.0642803,
similarity=23.601694]
3: 21148 Netherlands
    [pem_pos=0, final_scores=0.46147877,
final_local_scores=0.505355, pem=-0.7943452,
similarity=23.490288]
4: 1107824 RKC_Waalwijk
    [pem_pos=0, final_scores=0.599422,
final_local_scores=0.5613537, pem=0.0, similarity=20.841627]
5: 834275 Willem_II_(football_club)
    [pem_pos=0, final_scores=0.78656465,
final_local_scores=0.75588024, pem=0.0, similarity=26.646687]
6: 1336325 Fortuna_Sittard
    [pem_pos=0, final_scores=0.73174983,
final_local_scores=0.7083622, pem=0.0, similarity=25.228655]
7: 505173 NAC_Breda
    [pem_pos=0, final_scores=0.6677082,
final_local_scores=0.6330624, pem=0.0, similarity=22.981558]
8: 79818 Sparta_Rotterdam
    [pem_pos=0, final_scores=0.5715154,
final_local_scores=0.53133434, pem=0.0, similarity=19.94579]
9: 745757 SC_Herenveen
    [pem_pos=1, final_scores=0.66199565,
final_local_scores=0.62301534, pem=-0.6351416,
similarity=26.13573]
10: 2273 AFC_Ajax
    [pem_pos=0, final_scores=0.67540234,
final_local_scores=0.6375127, pem=0.0, similarity=23.114363]
11: 1336325 Fortuna_Sittard
    [pem_pos=0, final_scores=0.736263,
final_local_scores=0.7132922, pem=0.0, similarity=25.375774]
12: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.47954065,
final_local_scores=0.41969055, pem=0.0, similarity=16.614117]
13: 800033 FC_Twente
    [pem_pos=0, final_scores=0.5270919,
final_local_scores=0.4753909, pem=0.0, similarity=18.276325]
14: 834256 Vitesse
    [pem_pos=0, final_scores=0.5688425,
final_local_scores=0.5257067, pem=0.0, similarity=19.77785]
```

```

15: 745757 SC_Herenveen
    [pem_pos=1, final_scores=0.6280584,
final_local_scores=0.585944, pem=-0.6351416,
similarity=25.029448]
16: 505173 NAC_Breda
    [pem_pos=0, final_scores=0.67559415,
final_local_scores=0.6416766, pem=0.0, similarity=23.238623]
17: 2273 AFC_Ajax
    [pem_pos=0, final_scores=0.620982,
final_local_scores=0.5780667, pem=0.0, similarity=21.340376]
18: 834226 FC_Utrecht
    [pem_pos=2, final_scores=0.67584985,
final_local_scores=0.64273304, pem=-2.2398934,
similarity=35.451035]
19: 47707 Feyenoord
    [pem_pos=0, final_scores=0.59179217,
final_local_scores=0.54151803, pem=0.0, similarity=20.249691]
20: 834196 Roda_JC_Kerkrade
    [pem_pos=0, final_scores=0.6156963,
final_local_scores=0.58524144, pem=0.0, similarity=21.554485]
21: 2202202 FC_Volendam
    [pem_pos=1, final_scores=0.69020325,
final_local_scores=0.6714932, pem=-0.8958714,
similarity=29.000296]
22: 1103112 FC_Groningen
    [pem_pos=2, final_scores=0.7748347,
final_local_scores=0.7437095, pem=-1.9677447,
similarity=36.98439]
23: 1107824 RKC_Waalwijk
    [pem_pos=0, final_scores=0.56032664,
final_local_scores=0.51864797, pem=0.0, similarity=19.567204]
24: 79818 Sparta_Rotterdam
    [pem_pos=0, final_scores=0.6045262,
final_local_scores=0.5673936, pem=0.0, similarity=21.02187]
25: 834275 Willem_II_(football_club)
    [pem_pos=0, final_scores=0.7401858,
final_local_scores=0.7052184, pem=0.0, similarity=25.134836]
26: 2289 AZ_Alkmaar
    [pem_pos=0, final_scores=0.46915585,
final_local_scores=0.40899783, pem=0.0, similarity=16.295025]
27: 834154 N.E.C._(football_club)
    [pem_pos=0, final_scores=0.6075439,
final_local_scores=0.56491363, pem=0.0, similarity=20.947863]

```

Sample: b'273_SOCCER'

SOCCER - SUMMARIES OF [1 1 GERMAN 1] FIRST DIVISION
MATCHES . [2 2 BONN 2] 1996-08-24 Summaries of [3 3 German 3]
first division matches played on Saturday : [4 4 Bochum 4] 1
(1 Jack 66th minute) [5 5 Arminia Bielefeld 5] 1 (Molata
59th) . Halftime 0-0 . Attendance 25,000 [6 6 Borussia Moench
hengladbach 6] 1 ([7 7 Andersson 7] 22nd) [8 8 Karlsruhe 8]
3 (Haessler 33rd , 2 Dundee 45th , 3 Keller 90th) . 1-2 .

20,000 [9 9 Stuttgart 9] 2 (Balakow 50th , [10 10 Bobic 10]
61st) [11 11 Werder Bremen 11] 1 ([12 12 Votava 12]
68th) . 0-0 . 32,000 [13 13 1860 Munich 13] 1 ([14 14 Schwa
bl 14] 38th) [15 15 Borussia Dortmund 15] 3 ([16 16 Zorc 16]
59th-pen , Moeller 73rd , [17 17 Heinrich 17] 90th) . 1-0 .
50,000 [18 18 Bayer Leverkusen 18] 0 [19 19 Fortuna Duesseldo
rf 19] 1 ([20 20 Seeliger 20] 47th) . 0-0 . 18,000 [21 21 Fr
eiburg 21] 1 ([22 22 Zeyer 22] 52nd) [23 23 Cologne 23] 3
(Gaißmayer 9th , [24 24 Polster 24] 86th , 90th) . 0-1 .
22,500

1: 11867 Germany
2: 3295 Bonn
3: 11867 Germany
4: 513623 VfL_Bochem
5: 1145228 Arminia_Bielefeld
6: 263466 Borussia_Mönchengladbach
7: 740855 Patrik_Andersson
8: 1734592 Karlsruher_SC
9: 737608 VfB_Stuttgart
10: 2194769 Fredi_Bobic
11: 7598082 SV_Werder_Bremen
12: 12266793 Miroslav_Votava
13: 1007441 TSV_1860_München
14: 20976647 Manfred_Schwabl
15: 331715 Borussia_Dortmund
16: 9901734 Michael_Zorc
17: 5570243 Jörg_Heinrich
18: 502005 Bayer_04_Leverkusen
19: 1949688 Fortuna_Düsseldorf
20: 16089537 Thomas_Seeliger
21: 1438835 SC_Freiburg
22: 21062203 Andreas_Zeyer
23: 497888 1._FC_Köln
24: 1518363 Toni_Polster

1: 19577963 Mathias_Jack
[winner: 2230996 Jack_Shephard, pem_pos=0,
final_scores=-0.5568889, final_local_scores=-0.43380284,
pem=-2.1402142, similarity=2.782989]
RECALL MISS: 19577963 Mathias_Jack]

2: 3373736 Sean_Dundee
[winner: 306504 Dundee_F.C., pem_pos=1,
final_scores=-0.13222653, final_local_scores=-0.15188813,
pem=-1.3661808, similarity=6.986572]
RECALL MISS: 3373736 Sean_Dundee]

3: 10634409 Marc_Keller
[winner: 895763 Kasey_Keller, pem_pos=3,
final_scores=-0.3323447, final_local_scores=-0.3499999,
pem=-2.8574977, similarity=9.184535]
RECALL MISS: 10634409 Marc_Keller]

1: 11867 Germany
[pem_pos=1, final_scores=0.7747875,
final_local_scores=0.8487753, pem=-1.0465823,

```
similarity=35.11034]
2: 3295 Bonn
    [pem_pos=0, final_scores=0.4053076,
final_local_scores=0.47136402, pem=0.0, similarity=18.156155]
3: 11867 Germany
    [pem_pos=1, final_scores=0.6621055,
final_local_scores=0.7256873, pem=-1.0465823,
similarity=31.437147]
4: 513623 VfL_Bochum
    [pem_pos=1, final_scores=0.6737001,
final_local_scores=0.6508191, pem=-1.509258,
similarity=31.719032]
5: 1145228 Arminia_Bielefeld
    [pem_pos=0, final_scores=0.53564006,
final_local_scores=0.5006415, pem=0.0, similarity=19.029854]
6: 263466 Borussia_Mönchengladbach
    [pem_pos=0, final_scores=0.51099783,
final_local_scores=0.4639144, pem=0.0, similarity=17.933844]
7: 740855 Patrik_Andersson
    [pem_pos=13, final_scores=0.34280592,
final_local_scores=0.37237656, pem=-3.2582102,
similarity=32.920826]
8: 1734592 Karlsruher_SC
    [pem_pos=1, final_scores=0.7588635,
final_local_scores=0.73687655, pem=-2.2263577,
similarity=38.18686]
9: 737608 VfB_Stuttgart
    [pem_pos=1, final_scores=0.6417374,
final_local_scores=0.6044968, pem=-2.1952448,
similarity=34.067184]
10: 2194769 Fredi_Bobic
    [pem_pos=0, final_scores=0.5869978,
final_local_scores=0.6454624, pem=0.0, similarity=23.351599]
11: 7598082 SV_Werder_Bremen
    [pem_pos=0, final_scores=0.6402243,
final_local_scores=0.6054889, pem=0.0, similarity=22.158709]
12: 12266793 Miroslav_Votava
    [pem_pos=0, final_scores=0.6075847,
final_local_scores=0.6581438, pem=0.0, similarity=23.730038]
13: 1007441 TSV_1860_München
    [pem_pos=0, final_scores=0.4565249,
final_local_scores=0.41059005, pem=0.0, similarity=16.342539]
14: 20976647 Manfred_Schwabl
    [pem_pos=0, final_scores=0.4771927,
final_local_scores=0.5627246, pem=0.0, similarity=20.882538]
15: 331715 Borussia_Dortmund
    [pem_pos=0, final_scores=0.5261262,
final_local_scores=0.47919434, pem=0.0, similarity=18.389828]
16: 9901734 Michael_Zorc
    [pem_pos=0, final_scores=0.5345483,
final_local_scores=0.5716608, pem=0.0, similarity=21.149212]
17: 5570243 Jörg_Heinrich
```

```

[pem_pos=1, final_scores=0.9163564,
final_local_scores=0.9841097, pem=-2.0650008,
similarity=44.687305]
18: 502005 Bayer_04_Leverkusen
[pem_pos=0, final_scores=0.5147081,
final_local_scores=0.4651208, pem=0.0, similarity=17.969845]
19: 1949688 Fortuna_Düsseldorf
[pem_pos=0, final_scores=0.6026251,
final_local_scores=0.57536674, pem=0.0, similarity=21.259804]
20: 16089537 Thomas_Seeliger
[pem_pos=3, final_scores=0.5148609,
final_local_scores=0.6154283, pem=-1.0986123,
similarity=28.429747]
21: 1438835 SC_Freiburg
[pem_pos=2, final_scores=0.8316353,
final_local_scores=0.8333854, pem=-1.7450409,
similarity=38.449398]
22: 21062203 Andreas_Zeyer
[pem_pos=2, final_scores=0.30443913,
final_local_scores=0.4072343, pem=-1.0986123,
similarity=22.216822]
23: 497888 1._FC_Köln
[pem_pos=10, final_scores=0.7428091,
final_local_scores=0.7162204, pem=-3.09805,
similarity=42.310837]
24: 1518363 Toni_Polster
[pem_pos=0, final_scores=0.55197674,
final_local_scores=0.5747404, pem=0.0, similarity=21.241114]

```

Sample: b'274_SOCCER'

SOCCER - RESULTS OF [1 1 GERMAN 1] FIRST DIVISION MATCHES . [2 2 BONN 2] 1996-08-24 Results of [3 3 German 3] first division soccer matches played on Saturday : [4 4 Bochum 4] 1 [5 5 Arminia Bielefeld 5] 1 [6 6 Borussia Moenchengladbach 6] 1 [7 7 Karlsruhe 7] 3 [8 8 Stuttgart 8] 2 [9 9 Werder Bremen 9] 1 [10 10 1860 Munich 10] 1 [11 11 Borussia Dortmund 11] 3 [12 12 Bayer Leverkusen 12] 0 [13 13 Fortuna Duesseldorf 13] 1 [14 14 Freiburg 14] 1 [15 15 Cologne 15] 3 Played on Saturday : [16 16 St Pauli 16] 4 [17 17 Schalke 17] 4 [18 18 Hansa Rostock 18] 0 [19 19 Hamburg 19] 1 [20 20 Bundesliga 20] standings after Saturday 's games (tabulate under played , won , drawn , lost , goals for , goals against , points) : [21 21 Cologne 21] 3 3 0 0 7 1 9 [22 22 VfB Stuttgart 22] 2 2 0 0 6 1 6 [23 23 Borussia Dortmund 23] 3 2 0 1 9 5 6 [24 24 Hamburg 24] 3 2 0 1 7 3 6 [25 25 Bayer Leverkusen 25] 3 2 0 1 7 4 6 [26 26 VfL Bochum 26] 3 1 2 0 3 2 5 [27 27 Karlsruhe 27] 2 1 1 0 5 3 4 [28 28 Bayern Munich 28] 2 1 1 0 3 2 4 [29 29 St Pauli 29] 3 1 1 1 7 7 4 [30 30 1860 Munich 30] 3 1 0 2 3 5 3 [31 31 Freiburg 31] 3 1 0 2 5 10 3 [32 32 Fortuna Duesseldorf 32] 3 1 0 2 1 7 3 [33 33 Hansa Rostock 33] 3 0 2 1 3 4 2 [34 34 Arminia Bielefeld 34] 3 0 2 1 2 3 2 [35 35 Borussia Moenchengladbach 35] 3 0 2 1 1 3 2 [36 36 Schalke 36] 3 0 2 1 4 8 2 [37]

```
37 Werder Bremen 37] 3 0 1 2 4 6 1 [38 38 MSV Duisburg 38] 2
0 0 2 1 4 0
1: 11867 Germany
2: 3295 Bonn
3: 11867 Germany
4: 513623 VfL_Bochum
5: 1145228 Arminia_Bielefeld
6: 263466 Borussia_Mönchengladbach
7: 1734592 Karlsruher_SC
8: 737608 VfB_Stuttgart
9: 7598082 SV_Werder_Bremen
10: 1007441 TSV_1860_München
11: 331715 Borussia_Dortmund
12: 502005 Bayer_04_Leverkusen
13: 1949688 Fortuna_Düsseldorf
14: 1438835 SC_Freiburg
15: 497888 1._FC_Köln
16: 881244 FC_St._Pauli
17: 410348 FC_Schalke_04
18: 599978 F.C._Hansa_Rostock
19: 535879 Hamburger_SV
20: 686142 Bundesliga
21: 497888 1._FC_Köln
22: 737608 VfB_Stuttgart
23: 331715 Borussia_Dortmund
24: 535879 Hamburger_SV
25: 502005 Bayer_04_Leverkusen
26: 513623 VfL_Bochum
27: 1734592 Karlsruher_SC
28: 172326 FC_Bayern_Münich
29: 881244 FC_St._Pauli
30: 1007441 TSV_1860_München
31: 1438835 SC_Freiburg
32: 1949688 Fortuna_Düsseldorf
33: 599978 F.C._Hansa_Rostock
34: 1145228 Arminia_Bielefeld
35: 263466 Borussia_Mönchengladbach
36: 410348 FC_Schalke_04
37: 7598082 SV_Werder_Bremen
38: 1586579 MSV_Duisburg
1: 11867 Germany
    [pem_pos=1, final_scores=0.7679631,
final_local_scores=0.8547228, pem=-1.0465823,
similarity=35.287823]
2: 3295 Bonn
    [pem_pos=0, final_scores=0.39073807,
final_local_scores=0.4654609, pem=0.0, similarity=17.979994]
3: 11867 Germany
    [pem_pos=1, final_scores=0.7294068,
final_local_scores=0.8126059, pem=-1.0465823,
similarity=34.03097]
4: 513623 VfL_Bochum
```

```
[pem_pos=1, final_scores=0.66342336,
final_local_scores=0.6258288, pem=-1.509258,
similarity=30.97327]
5: 1145228 Arminia_Bielefeld
    [pem_pos=0, final_scores=0.54445714,
final_local_scores=0.49639964, pem=0.0, similarity=18.903269]
6: 263466 Borussia_Mönchengladbach
    [pem_pos=0, final_scores=0.5245216,
final_local_scores=0.46770597, pem=0.0, similarity=18.046991]
7: 1734592 Karlsruher_SC
    [pem_pos=1, final_scores=0.7048447,
final_local_scores=0.6686953, pem=-2.2263577,
similarity=36.152195]
8: 737608 VfB_Stuttgart
    [pem_pos=1, final_scores=0.6954053,
final_local_scores=0.6556626, pem=-2.1952448,
similarity=35.594074]
9: 7598082 SV_Werder_Bremen
    [pem_pos=0, final_scores=0.6082393,
final_local_scores=0.5598292, pem=0.0, similarity=20.796133]
10: 1007441 TSV_1860_München
    [pem_pos=0, final_scores=0.4761719,
final_local_scores=0.42234355, pem=0.0, similarity=16.693287]
11: 331715 Borussia_Dortmund
    [pem_pos=0, final_scores=0.5367624,
final_local_scores=0.4802962, pem=0.0, similarity=18.422709]
12: 502005 Bayer_04_Leverkusen
    [pem_pos=0, final_scores=0.56175464,
final_local_scores=0.5071464, pem=0.0, similarity=19.223974]
13: 1949688 Fortuna_Düsseldorf
    [pem_pos=0, final_scores=0.6238095,
final_local_scores=0.5837511, pem=0.0, similarity=21.51001]
14: 1438835 SC_Freiburg
    [pem_pos=2, final_scores=0.83584315,
final_local_scores=0.8284473, pem=-1.7450409,
similarity=38.302036]
15: 497888 1._FC_Köln
    [pem_pos=10, final_scores=0.6752855,
final_local_scores=0.62986815, pem=-3.09805,
similarity=39.733913]
16: 881244 FC_St._Pauli
    [pem_pos=1, final_scores=0.5781482,
final_local_scores=0.54195005, pem=-0.11686432,
similarity=20.89811]
17: 410348 FC_Schalke_04
    [pem_pos=0, final_scores=0.54526895,
final_local_scores=0.48782027, pem=0.0, similarity=18.647243]
18: 599978 F.C._Hansa_Rostock
    [pem_pos=0, final_scores=0.5761612,
final_local_scores=0.53693235, pem=0.0, similarity=20.112844]
19: 535879 Hamburger_SV
    [pem_pos=1, final_scores=0.86149067,
```

```
final_local_scores=0.8331101, pem=-2.820485,
similarity=44.289616]
20: 686142 Bundesliga
    [pem_pos=0, final_scores=0.38107193,
final_local_scores=0.31138968, pem=-0.08715299,
similarity=13.856154]
21: 497888 1._FC_Köln
    [pem_pos=10, final_scores=0.710079,
final_local_scores=0.6678748, pem=-3.09805,
similarity=40.868107]
22: 737608 VfB_Stuttgart
    [pem_pos=0, final_scores=0.44728643,
final_local_scores=0.38463044, pem=0.0, similarity=15.567853]
23: 331715 Borussia_Dortmund
    [pem_pos=0, final_scores=0.45934868,
final_local_scores=0.39573342, pem=0.0, similarity=15.899188]
24: 535879 Hamburger_SV
    [pem_pos=1, final_scores=0.76614934,
final_local_scores=0.7289641, pem=-2.820485,
similarity=41.181694]
25: 502005 Bayer_04_Leverkusen
    [pem_pos=0, final_scores=0.50893,
final_local_scores=0.44944358, pem=0.0, similarity=17.502007]
26: 513623 VfL_Bochum
    [pem_pos=0, final_scores=0.55817246,
final_local_scores=0.51085824, pem=0.0, similarity=19.334742]
27: 1734592 Karlsruher_SC
    [pem_pos=1, final_scores=0.67053807,
final_local_scores=0.6312205, pem=-2.2263577,
similarity=35.033875]
28: 172326 FC_Bayern_Munich
    [pem_pos=0, final_scores=0.4543864,
final_local_scores=0.39462638, pem=0.0, similarity=15.866151]
29: 881244 FC_St._Pauli
    [pem_pos=1, final_scores=0.5593141,
final_local_scores=0.5213766, pem=-0.11686432,
similarity=20.284157]
30: 1007441 TSV_1860_München
    [pem_pos=0, final_scores=0.46951407,
final_local_scores=0.4150709, pem=0.0, similarity=16.476255]
31: 1438835 SC_Freiburg
    [pem_pos=2, final_scores=0.83139855,
final_local_scores=0.82359225, pem=-1.7450409,
similarity=38.15715]
32: 1949688 Fortuna_Düsseldorf
    [pem_pos=0, final_scores=0.63226056,
final_local_scores=0.59298253, pem=0.0, similarity=21.785494]
33: 599978 F.C._Hansa_Rostock
    [pem_pos=0, final_scores=0.555435,
final_local_scores=0.5142922, pem=0.0, similarity=19.437218]
34: 1145228 Arminia_Bielefeld
    [pem_pos=0, final_scores=0.5158559,
```

```

final_local_scores=0.46515715, pem=0.0, similarity=17.97093]
35: 263466 Borussia_Mönchengladbach
    [pem_pos=0, final_scores=0.53303885,
final_local_scores=0.47700977, pem=0.0, similarity=18.324636]
36: 410348 FC_Schalke_04
    [pem_pos=0, final_scores=0.5471323,
final_local_scores=0.4898557, pem=0.0, similarity=18.707985]
37: 7598082 SV_Werder_Bremen
    [pem_pos=0, final_scores=0.6183627,
final_local_scores=0.5708875, pem=0.0, similarity=21.126135]
38: 1586579 MSV_Duisburg
    [pem_pos=0, final_scores=0.54160434,
final_local_scores=0.49752963, pem=0.0, similarity=18.936989]

```

Sample: b'275_SOCCER'

```

SOCCER - [1 1 AUSTRIA 1] FIRST DIVISION RESULTS /
STANDINGS . [2 2 VIENNA 2] 1996-08-24 Results of [3 3 Austria
3] first division soccer matches played on Saturday : [4 4 Ra
pid Vienna 4] 0 [5 5 FC Linz 5] 0 1 GAK 2 [6 6 Austria Vienna
6] 2 Admira / Wacker 0 [7 7 Sturm Graz 7] 3 [8 8 Linzer ASK 8]
1 [9 9 FC Tirol Innsbruck 9] 3 Standings ( tabulated under
played , won , drawn , lost , goals for , against ,
points ) : [10 10 FC Tirol Innsbruck 10] 6 4 2 0 13 5 14 [11 1
1 Austria Vienna 11] 6 4 2 0 9 5 14 [12 12 SV Salzburg 12] 5
3 2 0 4 1 11 [13 13 Sturm Graz 13] 6 2 3 1 8 5 9 2 GAK 6 1 3
2 8 10 6 [14 14 Rapid Wien 14] 5 0 5 0 3 3 5 [15 15 SV Ried 15
] 5 1 1 3 6 5 4 [16 16 Linzer ASK 16] 5 0 3 2 4 8 3 Admira /
Wacker 6 0 3 3 5 10 3 [17 17 FC Linz 17] 6 0 2 4 1 9 2

```

1: 26964606 Austria

2: 55866 Vienna

3: 26964606 Austria

4: 656932 SK_Rapid_Wien

5: 5644199 FC_Linz

6: 895568 FK_Austria_Wien

7: 1595328 SK_Sturm_Graz

8: 5644175 LASK_Linz

9: 5599205 FC_Tirol_Innsbruck

10: 5599205 FC_Tirol_Innsbruck

11: 895568 FK_Austria_Wien

12: 2163477 FC_Red_Bull_Salzburg

13: 1595328 SK_Sturm_Graz

14: 656932 SK_Rapid_Wien

15: 3004699 SV_Ried

16: 5644175 LASK_Linz

17: 5644199 FC_Linz

1: 887879 Grazer_AK

```

[winner: 3968641 Nickelodeon_Toys, pem_pos=0,
final_scores=-0.4240406, final_local_scores=-0.2724986,
pem=0.0, similarity=-4.0421886]
RECALL MISS: 887879 Grazer_AK]

```

2: 887879 Grazer_AK

```

[winner: 3968641 Nickelodeon_Toys, pem_pos=0,
```

```
final_scores=-0.37447435, final_local_scores=-0.21835503,
pem=0.0, similarity=-2.426437]
    RECALL MISS: 887879 Grazer_AK]
1: 26964606 Austria
    [pem_pos=0, final_scores=0.6163431,
final_local_scores=0.67165524, pem=-0.25262418,
similarity=25.507055]
2: 55866 Vienna
    [pem_pos=0, final_scores=0.50331867,
final_local_scores=0.5664388, pem=-0.12517071,
similarity=21.674074]
3: 26964606 Austria
    [pem_pos=0, final_scores=0.49818033,
final_local_scores=0.54258037, pem=-0.25262418,
similarity=21.655203]
4: 656932 SK_Rapid_Wien
    [pem_pos=0, final_scores=0.5107743,
final_local_scores=0.46866095, pem=0.0, similarity=18.07549]
5: 5644199 FC_Linz
    [pem_pos=1, final_scores=0.3361464,
final_local_scores=0.34856832, pem=-0.50743,
similarity=17.251171]
6: 895568 FK_Austria_Wien
    [pem_pos=0, final_scores=0.5704939,
final_local_scores=0.5268081, pem=0.0, similarity=19.810717]
7: 1595328 SK_Sturm_Graz
    [pem_pos=0, final_scores=0.4710794,
final_local_scores=0.43120873, pem=0.0, similarity=16.957842]
8: 5644175 LASK_Linz
    [pem_pos=0, final_scores=0.56269664,
final_local_scores=0.5305927, pem=0.0, similarity=19.923656]
9: 5599205 FC_Tirol_Innsbruck
    [pem_pos=0, final_scores=0.41745204,
final_local_scores=0.41606808, pem=0.0, similarity=16.506014]
10: 5599205 FC_Tirol_Innsbruck
    [pem_pos=0, final_scores=0.43961024,
final_local_scores=0.44027257, pem=0.0, similarity=17.228325]
11: 895568 FK_Austria_Wien
    [pem_pos=0, final_scores=0.57393247,
final_local_scores=0.5305642, pem=0.0, similarity=19.922806]
12: 2163477 FC_Red_Bull_Salzburg
    [pem_pos=0, final_scores=0.41746312,
final_local_scores=0.38141286, pem=0.0, similarity=15.471833]
13: 1595328 SK_Sturm_Graz
    [pem_pos=0, final_scores=0.4593882,
final_local_scores=0.41843784, pem=0.0, similarity=16.576733]
14: 656932 SK_Rapid_Wien
    [pem_pos=0, final_scores=0.4030823,
final_local_scores=0.35102376, pem=0.0, similarity=14.564962]
15: 3004699 SV_Ried
    [pem_pos=0, final_scores=0.5386177,
final_local_scores=0.53681606, pem=0.0, similarity=20.109375]
```

```
16: 5644175 LASK_Linz
    [pem_pos=0, final_scores=0.5767885,
final_local_scores=0.5459858, pem=0.0, similarity=20.383018]
17: 5644199 FC_Linz
    [pem_pos=1, final_scores=0.33555883,
final_local_scores=0.34792644, pem=-0.50743,
similarity=17.232016]
```

Sample: b'276_CRICKET'

CRICKET - RAIN BRINGS PREMATURE END TO [1 1 SRI LANKA 1]
MATCH . [2 2 COLOMBO 2] 1996-08-24 The one-day match between [3 3 Sri Lanka 3] and a [4 4 World XI 4] was abandoned on Saturday because of rain . Scores : [5 5 World XI 5] 102-0 (M. Waugh 39 not out , S. Tendulkar 56 not out) off 21.4 overs v [6 6 Sri Lanka 6] .

```
1: 26750 Sri_Lanka
2: 56636 Colombo
3: 397246 Sri_Lanka_national_cricket_team
4: 23538966 World_Series_Cricket_World_XI
5: 23538966 World_Series_Cricket_World_XI
6: 397246 Sri_Lanka_national_cricket_team
1: 26750 Sri_Lanka
    [pem_pos=0, final_scores=0.3753184,
final_local_scores=0.4254893, pem=-0.3027737,
similarity=18.433693]
2: 56636 Colombo
    [pem_pos=0, final_scores=0.54125637,
final_local_scores=0.5889768, pem=-0.19393882,
similarity=22.720625]
3: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.71256906,
final_local_scores=0.7164603, pem=-2.2667584,
similarity=37.797302]
4: 23538966 World_Series_Cricket_World_XI
    [pem_pos=0, final_scores=0.7093303,
final_local_scores=0.78092295, pem=0.0, similarity=27.39401]
5: 23538966 World_Series_Cricket_World_XI
    [pem_pos=0, final_scores=0.6620166,
final_local_scores=0.7292398, pem=0.0, similarity=25.851685]
6: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.58315986,
final_local_scores=0.57510036, pem=-2.2667584,
similarity=33.57884]
```

Sample: b'277_British'

[1 1 British 1] hostage in [2 2 Chechnya 2] describes ordeal . [3 3 LONDON 3] 1996-08-24 A [4 4 British 4] aid worker , held hostage in [5 5 Chechnya 5] for nearly four weeks , said on Saturday a cocked [6 6 Kalashnikov 6] had been thrust into his mouth at one point during his ordeal . Michael Penrose , a 23-year-old worker with the group [7 7 Action Against Hunger 7] , described his ordeal to a news

conference when he arrived back in [8 8 Britain 8] from [9 9 M oscow 9] . " The worst period of physical manhandling was during that time when we were beaten with [10 10 Kalashnikovs 10] and at one point I had a [11 11 Kalashnikov 11] held to the back of my throat -- cocked , " he said . " For the first period we were held in a small room with no bed or anything . We had very little food and sometimes went two or three days without eating . " Gunmen seized Penrose , who comes from [12 12 Swerford 12] in southern [13 13 England 13] , [14 14 Frenc hman 14] Frederic Malardeau and six other hostages from their car in [15 15 Grozny 15] , the capital of [16 16 Chechnya 16] , on July 27 . The assailants had demanded a ransom of 300,000 pounds (\$ 465,000) but no money was paid by the charity when they were released on Wednesday . The hostages were held in a house in or near [17 17 Grozny 17] which was bombarded regularly . " During the last 15 days of being held , the fighting in [18 18 Grozny 18] was very close . At first it was street fighting outside the house and then we came under very heavy shelling and bombardment from conventional weapons like tanks , artillery and grenade launchers , " he said . Penrose had been working for the charity which provides food to civilians for only a few weeks before he was captured . When asked if he would return to the mountainous region where rebels have been fighting [19 19 Rus sian 19] troops for full independence , Penrose said : " Not for the time being . I do n't think it 's safe for me . But maybe in the future , depending on the circumstances . "

1: 31717 United_Kingdom
2: 6095 Chechnya
3: 17867 London
4: 31717 United_Kingdom
5: 6095 Chechnya
6: 1348 AK-47
7: 2448 Action_Against_Hunger
8: 31717 United_Kingdom
9: 19004 Moscow
10: 1348 AK-47
11: 1348 AK-47
12: 21592637 Swerford
13: 9316 England
14: 5843419 France
15: 140790 Grozny
16: 6095 Chechnya
17: 140790 Grozny
18: 140790 Grozny
19: 25391 Russia
1: 31717 United_Kingdom
[pem_pos=0, final_scores=0.4243148,
final_local_scores=0.40468705, pem=-0.48936236,
similarity=18.82761]
2: 6095 Chechnya
[pem_pos=0, final_scores=0.5482311,

```
final_local_scores=0.5334143, pem=0.0, similarity=20.00786]
3: 17867 London
    [pem_pos=0, final_scores=0.6224257,
final_local_scores=0.64074343, pem=-0.09086737,
similarity=23.704926]
4: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.48483092,
final_local_scores=0.4707917, pem=-0.48936236,
similarity=20.800304]
5: 6095 Chechnya
    [pem_pos=0, final_scores=0.5058808,
final_local_scores=0.487153, pem=0.0, similarity=18.62733]
6: 1348 AK-47
    [pem_pos=0, final_scores=0.2782997,
final_local_scores=0.3157064, pem=-0.13110125,
similarity=14.22397]
7: 2448 Action_Against_Hunger
    [pem_pos=0, final_scores=0.4173059,
final_local_scores=0.5313014, pem=0.0, similarity=19.944805]
8: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.47050327,
final_local_scores=0.455141, pem=-0.61114657,
similarity=20.995537]
9: 19004 Moscow
    [pem_pos=0, final_scores=0.75967807,
final_local_scores=0.76790285, pem=-0.106582455,
similarity=27.585077]
10: 1348 AK-47
    [pem_pos=0, final_scores=0.3386457,
final_local_scores=0.38162524, pem=0.0, similarity=15.47817]
11: 1348 AK-47
    [pem_pos=0, final_scores=0.3795035,
final_local_scores=0.42625618, pem=-0.13110125,
similarity=17.522997]
12: 21592637 Swerford
    [pem_pos=0, final_scores=0.50493664,
final_local_scores=0.6865909, pem=0.0, similarity=24.578955]
13: 9316 England
    [pem_pos=0, final_scores=0.50607455,
final_local_scores=0.5289135, pem=-0.19053856,
similarity=20.909725]
14: 5843419 France
    [pem_pos=1, final_scores=0.4905352,
final_local_scores=0.50455177, pem=-0.7184992,
similarity=23.053854]
15: 140790 Grozny
    [pem_pos=0, final_scores=0.5079847,
final_local_scores=0.5088956, pem=0.0, similarity=19.276173]
16: 6095 Chechnya
    [pem_pos=0, final_scores=0.5748849,
final_local_scores=0.5625295, pem=0.0, similarity=20.876715]
17: 140790 Grozny
```

```
[pem_pos=0, final_scores=0.5736451,
final_local_scores=0.5806196, pem=0.0, similarity=21.41656]
18: 140790 Grozny
    [pem_pos=0, final_scores=0.49738562,
final_local_scores=0.49731767, pem=0.0, similarity=18.930664]
19: 25391 Russia
    [pem_pos=1, final_scores=0.88098377,
final_local_scores=0.9230111, pem=-1.0515022,
similarity=37.35244]
```

Sample: b'278_Princess'

Princess [1 1 Diana 1] send message to [2 2 Mother Teresa 2]. [3 3 LONDON 3] 1996-08-24 [4 4 Britain 4] 's Princess [5 5 Diana 5] has sent a message to seriously ill [6 6 Mother Teresa 6], the nun to whom she has turned several times for spiritual guidance. [7 7 Diana 7] 's office said on Saturday the princess had sent a message to the Nobel Peace Prize-winning missionary as news broke this week of her battle against heart problems and malaria. A spokeswoman declined to release details of the message. [8 8 Diana 8] first met the Albanian-born missionary in [9 9 Rome 9] in 1992. She said afterwards that the meeting had fulfilled her "dearest wish" and the two women have met several times since. The princess, who has carved out a major role for herself as a helper of the sick and needy, is said to have turned to [10 10 Mother Teresa 10] for guidance as her marriage crumbled to heir to the [11 11 British 11] throne [12 12 Prince Charles 12]. The 85-year-old nun said in the past that she was praying for the couple, whose divorce is expected to become final next week. Doctors caring for [13 13 Mother Teresa 13] in a [14 14 Calcutta 14] hospital said on Saturday that her fever had fallen and her malaria was under control but she remained on a respirator in intensive care.

```
1: 24095 Diana,_Princess_of_Wales
2: 347105 Mother_Teresa
3: 17867 London
4: 31717 United_Kingdom
5: 24095 Diana,_Princess_of_Wales
6: 347105 Mother_Teresa
7: 24095 Diana,_Princess_of_Wales
8: 24095 Diana,_Princess_of_Wales
9: 25458 Rome
10: 347105 Mother_Teresa
11: 31717 United_Kingdom
12: 125248 Charles,_Prince_of_Wales
13: 347105 Mother_Teresa
14: 47905 Kolkata
1: 24095 Diana,_Princess_of_Wales
    [pem_pos=0, final_scores=1.0294898,
final_local_scores=1.0311754, pem=0.0, similarity=34.86204]
2: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.515565,
```

```
final_local_scores=0.5251631, pem=0.0, similarity=19.761627]
3: 17867 London
    [pem_pos=0, final_scores=0.63647884,
final_local_scores=0.6481908, pem=-0.09086737,
similarity=23.92717]
4: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.51263434,
final_local_scores=0.53272194, pem=-0.61114657,
similarity=23.310707]
5: 24095 Diana,_Princess_of_Wales
    [pem_pos=0, final_scores=1.037071,
final_local_scores=1.0394568, pem=0.0, similarity=35.10918]
6: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.62353665,
final_local_scores=0.6431058, pem=0.0, similarity=23.281273]
7: 24095 Diana,_Princess_of_Wales
    [pem_pos=0, final_scores=1.0875599,
final_local_scores=1.0946083, pem=0.0, similarity=36.755005]
8: 24095 Diana,_Princess_of_Wales
    [pem_pos=0, final_scores=1.1924202,
final_local_scores=1.2091522, pem=0.0, similarity=40.173225]
9: 25458 Rome
    [pem_pos=0, final_scores=0.50055254,
final_local_scores=0.5396928, pem=-0.2517334,
similarity=21.564188]
10: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.5929969,
final_local_scores=0.6097458, pem=0.0, similarity=22.285744]
11: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.41690594,
final_local_scores=0.42815322, pem=-0.48936236,
similarity=19.527887]
12: 125248 Charles,_Prince_of_Wales
    [pem_pos=0, final_scores=0.5534079,
final_local_scores=0.5271553, pem=-0.08443151,
similarity=20.28023]
13: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.57717246,
final_local_scores=0.5924599, pem=0.0, similarity=21.7699]
14: 47905 Kolkata
    [pem_pos=0, final_scores=0.3200435,
final_local_scores=0.34305042, pem=-0.02273471,
similarity=14.450656]
```

Sample: b'279_CRICKET'

```
CRICKET - [1 1 PAKISTAN 1] 339-4 V [2 2 ENGLAND 2] - close .
[3 3 Saeed Anwar 3] c [4 4 Croft 4] b [5 5 Cork 5] 176 Aamir
Sohail c [6 6 Cork 6] b [7 7 Croft 7] 46 [8 8 Ijaz Ahmed 8] c
1 Stewart b [9 9 Mullally 9] 61 [10 10 Inzamam-ul-Haq 10] c [1
1 11 Hussain 11] b [12 12 Mullally 12] 35 [13 13 Salim Malik 1
3] not out 2 [14 14 Asif Mujtaba 14] not out 1 Extras 18 Fall
of wicket - 1-106 2-239 3-334 4-334 To bat - [15 15 Wasim Akr
```

am 15] , [16 16 Moin Khan 16] , [17 17 Mushtaq Ahmed 17] , [18 18 Waqar Younis 18] , Mohammad Akam [1 2 England 1] 326 all out

1: 759295 Pakistan_national_cricket_team
2: 759125 England_cricket_team
3: 894437 Saeed_Anwar
4: 1160680 Robert_Croft
5: 2036840 Dominic_Cork
6: 2036840 Dominic_Cork
7: 1160680 Robert_Croft
8: 938863 Ijaz_Ahmed_(cricketer)
9: 2222543 Alan_Mullally
10: 896403 Inzamam-ul-Haq
11: 682925 Nasser_Hussain
12: 2222543 Alan_Mullally
13: 938861 Saleem_Malik
14: 3976606 Asif_Mujtaba
15: 501393 Wasim_Akram
16: 554132 Moin_Khan
17: 896405 Mushtaq_Ahmed
18: 894433 Waqar_Younis
1: 896414 Alec_Stewart
 [winner: 876610 Dave_Stewart_(baseball), pem_pos=14,
final_scores=-0.7341537, final_local_scores=-0.6798545,
pem=-4.102782, similarity=6.1130643]
 RECALL MISS: 896414 Alec_Stewart]

2: 9316 England
 [winner: 759125 England_cricket_team, pem_pos=2,
final_scores=0.2730853, final_local_scores=0.2145589,
pem=-3.3054674, similarity=28.468224]
 gt: 9316 England, pem_pos=0, final_scores=-0.054526493,
final_local_scores=0.012309462, pem=-0.19053856,
similarity=5.49323]

1: 759295 Pakistan_national_cricket_team
 [pem_pos=1, final_scores=0.76352245,
final_local_scores=0.7507933, pem=-2.7556365,
similarity=41.48046]

2: 759125 England_cricket_team
 [pem_pos=2, final_scores=0.6998758,
final_local_scores=0.67580193, pem=-3.3054674,
similarity=42.232635]

3: 894437 Saeed_Anwar
 [pem_pos=0, final_scores=0.74004465,
final_local_scores=0.747073, pem=0.0, similarity=26.383862]

4: 1160680 Robert_Croft
 [pem_pos=7, final_scores=0.9618264,
final_local_scores=1.0043082, pem=-3.16984,
similarity=51.298355]

5: 2036840 Dominic_Cork
 [pem_pos=8, final_scores=0.892844,
final_local_scores=0.9037264, pem=-3.606134,
similarity=50.66943]

```
6: 2036840 Dominic_Cork
    [pem_pos=8, final_scores=1.130728,
final_local_scores=1.1635785, pem=-3.606134,
similarity=58.423935]
7: 1160680 Robert_Croft
    [pem_pos=7, final_scores=0.9755979,
final_local_scores=1.0193515, pem=-3.16984,
similarity=51.747276]
8: 938863 Ijaz_Ahmed_(cricketer)
    [pem_pos=0, final_scores=0.61852664,
final_local_scores=0.6175463, pem=0.0, similarity=22.518526]
9: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.7582262,
final_local_scores=0.7834364, pem=-0.73930055,
similarity=31.48945]
10: 896403 Inzamam-ul-Haq
    [pem_pos=0, final_scores=1.0154898,
final_local_scores=1.0385506, pem=0.0, similarity=35.082134]
11: 682925 Nasser_Hussain
    [pem_pos=3, final_scores=0.6778176,
final_local_scores=0.68072873, pem=-2.447519,
similarity=37.714005]
12: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.81236917,
final_local_scores=0.84257936, pem=-0.73930055,
similarity=33.254395]
13: 938861 Saleem_Malik
    [pem_pos=0, final_scores=0.7379033,
final_local_scores=0.7666923, pem=0.0, similarity=26.96934]
14: 3976606 Asif_Mujtaba
    [pem_pos=0, final_scores=0.7699509,
final_local_scores=0.80874276, pem=0.0, similarity=28.22421]
15: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.7430431,
final_local_scores=0.73940796, pem=0.0, similarity=26.15512]
16: 554132 Moin_Khan
    [pem_pos=0, final_scores=0.70175415,
final_local_scores=0.74377406, pem=0.0, similarity=26.285414]
17: 896405 Mushtaq_Ahmed
    [pem_pos=0, final_scores=0.6191283,
final_local_scores=0.61892563, pem=0.0, similarity=22.559689]
18: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.70947057,
final_local_scores=0.71025705, pem=0.0, similarity=25.2852]
1: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.2730853,
final_local_scores=0.2145589, pem=-3.3054674,
similarity=28.468224]
```

Sample: b'280_Soccer'

Soccer - Nijmeh beat Nasr 1-0 . [1 1 TRIPOLI 1] , [2 2 Lebanon 2] 1996-08-24 Nijmeh of [3 3 Lebanon 3] beat Nasr of [4 4 S

```
audi Arabia 4] 1-0 ( halftime 1-0 ) in their [5 5 Asian 5]
club championship second round first leg tie on Saturday .
Scorer : Issa Alloush ( 45th minute ) . Attendance : 10,000 .
1: 268796 Tripoli,_Lebanon
2: 17771 Lebanon
3: 17771 Lebanon
4: 349303 Saudi_Arabia
5: 689 Asia
1: 268796 Tripoli,_Lebanon
    [pem_pos=1, final_scores=0.2096598,
final_local_scores=0.21116996, pem=-1.8286858,
similarity=20.336123]
2: 17771 Lebanon
    [pem_pos=0, final_scores=0.5173482,
final_local_scores=0.46123022, pem=-0.4097839,
similarity=20.08221]
3: 17771 Lebanon
    [pem_pos=0, final_scores=0.62540543,
final_local_scores=0.57926637, pem=-0.4097839,
similarity=23.604647]
4: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.4520741,
final_local_scores=0.4128632, pem=-0.25454172,
similarity=17.794611]
5: 689 Asia
    [pem_pos=1, final_scores=0.41955847,
final_local_scores=0.46420574, pem=-1.9623682,
similarity=28.614202]
```

Sample: b'281_Adventurers'

Adventurers start [1 1 Canadian 1] wilderness race . [2 2 PE MBERTON 2] , [3 3 British Columbia 3] 1996-08-24 About 350 adventurers from nine countries set out on Saturday to climb , raft , bike and run in a 323-mile (517-km) endurance race through the [4 4 Canadian 4] wilderness . The event , called the Eco-Challenge , is part of a growing sport known as adventure racing in which competitors test their limits for days over a perilous wilderness course . " I 'm looking forward to this race . I think it will be more physically challenging and we 'll have to go up against more diverse situations due to the terrain , " said Dr. Michael Stroud , a veteran Eco-Challenge participant . The Eco-Challenge has been staged twice before -- in [5 5 Utah 5] and Maine last year -- and is modelled on similar races overseas . The 70 teams in this year 's race will will trek glaciers , climb mountains , whitewater raft , horseback ride , canoe and mountain bike along the grueling course . This year 's race , the route of which was keep a secret until Friday evening , is being held near [6 6 Pemberton 6] , [7 7 British Columbia 7] , about 100 miles (160 km) northeast of [8 8 Vancouver 8] . The area is filled with treacherous mountain peaks , ice fields and frigid waters .

Organisers expect about two-thirds of the participants to drop out or be disqualified before the finish . The hardy ones are expected to complete the course in about six days , with first-place finishers receiving \$ 10,000 in prize money . In the Eco-Challenge , competitors race in teams of five which must include both men and women . Team members must remain within 100 yards (metres) of each other at all times and finish together . With racers carrying about 40 pounds (18 kg) of gear on their backs , broken bones , sunstroke , dehydration and exhaustion are common .

```
1: 5042916 Canada
2: 3582046 Pemberton,_British_Columbia
3: 3392 British_Columbia
4: 5042916 Canada
5: 31716 Utah
6: 3582046 Pemberton,_British_Columbia
7: 3392 British_Columbia
8: 32706 Vancouver
1: 5042916 Canada
    [pem_pos=0, final_scores=0.6739823,
final_local_scores=0.6496612, pem=-0.41303828,
similarity=25.723064]
2: 3582046 Pemberton,_British_Columbia
    [pem_pos=0, final_scores=0.34575462,
final_local_scores=0.34086615, pem=-1.3705597,
similarity=21.715155]
3: 3392 British_Columbia
    [pem_pos=0, final_scores=0.485084,
final_local_scores=0.42472732, pem=-0.028628228,
similarity=16.920107]
4: 5042916 Canada
    [pem_pos=0, final_scores=0.64710134,
final_local_scores=0.6202978, pem=-0.41303828,
similarity=24.846802]
5: 31716 Utah
    [pem_pos=0, final_scores=0.53747946,
final_local_scores=0.5443483, pem=-0.2611261,
similarity=21.754196]
6: 3582046 Pemberton,_British_Columbia
    [pem_pos=0, final_scores=0.46015662,
final_local_scores=0.4658329, pem=-1.3705597,
similarity=25.444412]
7: 3392 British_Columbia
    [pem_pos=0, final_scores=0.5237506,
final_local_scores=0.46696472, pem=-0.028628228,
similarity=18.180555]
8: 32706 Vancouver
    [pem_pos=0, final_scores=0.5198291,
final_local_scores=0.5117434, pem=-0.12074051,
similarity=20.017763]
```

Sample: b'282_Cholera'

Cholera kills 21 in southern [1 1 Nigeria 1] . [2 2 LAGOS 2] 1996-05-28 An outbreak of cholera has killed 21 people in a week at Ubimini in oil-rich southern [3 3 Nigeria 3] , the News Agency of Nigeria reported on Saturday . The chairman of the local council , Damian Ejiohuo , said drugs had been rushed to the area to quell the disease and the community needed a safer source of drinking water to prevent future outbreaks . Epidemics are common in rural areas of [4 4 Nigeria 4] where piped water is not usually available .

1: 21383 Nigeria
2: 85232 Lagos
3: 21383 Nigeria
4: 21383 Nigeria
1: 21383 Nigeria
[pem_pos=0, final_scores=0.6127563,
final_local_scores=0.5501492, pem=-0.31455088,
similarity=22.21784]
2: 85232 Lagos
[pem_pos=0, final_scores=0.46668446,
final_local_scores=0.4618566, pem=-0.20534219,
similarity=18.989119]
3: 21383 Nigeria
[pem_pos=0, final_scores=0.56270003,
final_local_scores=0.4954703, pem=-0.31455088,
similarity=20.586111]
4: 21383 Nigeria
[pem_pos=0, final_scores=0.5211224,
final_local_scores=0.4500531, pem=-0.31455088,
similarity=19.230772]

Sample: b'283_Malawi'

[1 1 Malawi 1] 's ex-president [2 2 Banda 2] says he 's feeling well . [3 3 BLANTYRE 3] 1996-08-24 [4 4 Malawi 4] 's frail former president , Kamuzu Banda , said on Saturday in a rare public interview that he was feeling well despite his advanced years . " I feel all right and I eat everything that is put on the table . And that means I am all right , " he told reporters invited to his home . [5 5 Banda 5] , a vegetarian teetotaller believed to be 97 , walked unaided but supporting himself on a walking stick . He clutched a fly whisk which for a long time symbolised his obsession with power . His health was the subject of much recent speculation . [6 6 Malawi 6] 's undisputed ruler for three decades , he lost power in the first all-party elections in 1994 . He spent a year under house arrest and was tried but acquitted last year on charges of ordering the murder of four opponents in 1983 .

1: 19088 Malawi
2: 213567 Hastings_Banda
3: 642604 Blantyre
4: 19088 Malawi
5: 213567 Hastings_Banda

```
6: 19088 Malawi
1: 19088 Malawi
    [pem_pos=0, final_scores=0.5081453,
final_local_scores=0.4983859, pem=-0.27577603,
similarity=20.462255]
2: 213567 Hastings_Banda
    [pem_pos=0, final_scores=0.9035668,
final_local_scores=0.92525536, pem=0.0, similarity=31.70118]
3: 642604 Blantyre
    [pem_pos=0, final_scores=0.35813153,
final_local_scores=0.3744164, pem=0.0, similarity=15.263045]
4: 19088 Malawi
    [pem_pos=0, final_scores=0.5278869,
final_local_scores=0.5199507, pem=-0.27577603,
similarity=21.105791]
5: 213567 Hastings_Banda
    [pem_pos=0, final_scores=0.9602391,
final_local_scores=0.9871612, pem=0.0, similarity=33.548573]
6: 19088 Malawi
    [pem_pos=0, final_scores=0.47262555,
final_local_scores=0.45958602, pem=-0.27577603,
similarity=19.30439]
```

Sample: b'284_Zimbabwe'

[1 1 Zimbabwe 1] fires striking civil servants . Emelia Sithole [2 2 HARARE 2] 1996-08-24 The [3 3 Zimbabwean 3] government fired thousands of workers on Saturday for defying an order to end a strike which has crippled essential services and disrupted international and domestic flights . The [4 4 Public Service Commission 4] (PSC) said in a statement that the workers -- including nurses , junior doctors , mortuary attendants , customs officers and firefighters -- would be barred from entering their workplaces on Monday . " All civil servants who did not return to work at their normal working hours , and remained working for the full working day on 23 August 1996 , have been summarily dismissed ... with immediate effect , " it said in a statement . Union officials from the Public Service Association (PSA) were unavailable for comment . Public Service , Labour and Social Welfare Minister Florence Chitauro told state radio her ministry had already begun recruiting other people to replace the strikers , sub-contracting some of the work to private firms . The government had been threatening to fire the workers since the strike began on Tuesday , saying it was illegal . But the strikers ignored the threat and vowed to stay on the streets until their demands for wage rises of 30 to 60 percent were met . The stoppage has left essential services stretched with many hospitals handling only emergency cases under senior doctors with the help of army medical personnel and the [5 5 Red Cross 5] . It has also disrupted flights . Some internal services were cancelled , leaving tourists at the [6 6 Victor

ia Falls 6] resort stranded , and flights abroad were delayed . The PSA said 80 percent of the country 's 180,000 civil servants took part in the strike which is a rare challenge to President [7 7 Robert Mugabe 7] , who has been in power since independence from [8 8 Britain 8] in 1980 . Opposition parties , civic organisations and private-sector unions have expressed support for the action and denounced the government 's pay rises of up to eight percent for its workers . Civil servants earn on average Z\$ 1,000 (\$ 100) a month . They say their pay has not kept up at all with inflation , currently running at 22 percent .

1: 34399 Zimbabwe
2: 56680 Harare
3: 34399 Zimbabwe
4: 2509550 Public_utilities_commission
5: 5023832 International_Committee_of_the_Red_Cross
6: 1759823 Victoria_Falls,_Zimbabwe
7: 17380298 Robert_Mugabe
8: 31717 United_Kingdom
1: 34399 Zimbabwe
 [pem_pos=0, final_scores=0.44878107,
final_local_scores=0.45095986, pem=-0.36737105,
similarity=19.545076]
2: 56680 Harare
 [pem_pos=0, final_scores=0.38693935,
final_local_scores=0.3915907, pem=0.0, similarity=15.775561]
3: 34399 Zimbabwe
 [pem_pos=0, final_scores=0.4608047,
final_local_scores=0.4640938, pem=0.0, similarity=17.939198]
4: 2509550 Public_utilities_commission
 [pem_pos=3, final_scores=0.27045655,
final_local_scores=0.3965295, pem=-1.5661345,
similarity=24.439825]
5: 5023832 International_Committee_of_the_Red_Cross
 [pem_pos=1, final_scores=0.36810547,
final_local_scores=0.43571037, pem=-1.9555908,
similarity=27.726986]
6: 1759823 Victoria_Falls,_Zimbabwe
 [pem_pos=1, final_scores=0.36357045,
final_local_scores=0.4737153, pem=-0.6637333,
similarity=21.835808]
7: 17380298 Robert_Mugabe
 [pem_pos=0, final_scores=0.6147401,
final_local_scores=0.65237653, pem=0.0, similarity=23.55793]
8: 31717 United_Kingdom
 [pem_pos=0, final_scores=0.50405467,
final_local_scores=0.51892847, pem=-0.61114657,
similarity=22.899082]

Sample: b'285_Rwanda'
[1 1 Rwanda 1] says [2 2 Zaire 2] expels 28 [3 3 Rwandan 3]
refugees . [4 4 KIGALI 4] 1996-08-24 [5 5 Rwanda 5] said on

Saturday that [1 Zaire 1] had expelled 28 [6 6 Rwandan 6] [7 7 Hutu 7] refugees accused of being " trouble-makers " in camps in eastern [8 8 Zaire 8] . Captain Firmin Gatera , spokesman for the Tutsi-dominated [9 9 Rwandan 9] army , told [10 10 Reuters 10] in [11 11 Kigali 11] that 17 of the 28 refugees handed over on Friday from the Zairean town of [12 12 Goma 12] had been soldiers in the former [13 13 Hutu 13] army which fled to [14 14 Zaire 14] in 1994 after being defeated by [15 1 5 Tutsi 15] forces in [16 16 Rwanda 16] 's civil war . Zairean Prime Minister [17 17 Kengo wa Dondo 17] said on Thursday in a visit to [18 18 Rwanda 18] that his country would expell all the refugees back to [19 19 Rwanda 19] but he gave no timeframe . Zaire is home to 1.1 million [20 20 Rwanda 20] [21 21 Hutu 21] refugees who fled three months of civil war in 1994 . Many had taken part in the genocide that year of one million people , mostly [22 22 Tutsis 22] , and refuse to go home for fear of reprisal at the hands of the new Tutsi-dominated government in [23 23 Kigali 23] . Gatera said the refugees were handed over following a deal made at a meeting between the governor of [2 Zaire 2] 's north [24 24 Kivu 24] region and his counterpart in the [25 25 Rwandan 25] border town of [26 26 Gisenyi 26] . " After a meeting between the governor of north [27 27 Kivu 27] and the prefect of [28 2 8 Gisenyi 28] , 28 prisoners (refugees) were handed over to [29 29 Rwandan 29] authorities on Friday , " Gatera said . " Out of these 17 were former soldiers . These people are now in [30 30 Gisenyi 30] prison , " Gatera added .

1: 25645 Rwanda
2: 76762 Democratic_Republic_of_the_Congo
3: 25645 Rwanda
4: 56690 Kigali
5: 25645 Rwanda
6: 25645 Rwanda
7: 59537 Hutu
8: 76762 Democratic_Republic_of_the_Congo
9: 25645 Rwanda
10: 18998750 Reuters
11: 56690 Kigali
12: 1210483 Goma
13: 59537 Hutu
14: 76762 Democratic_Republic_of_the_Congo
15: 67050 Tutsi
16: 25645 Rwanda
17: 420869 Léon_Kengo
18: 25645 Rwanda
19: 25645 Rwanda
20: 25645 Rwanda
21: 59537 Hutu
22: 67050 Tutsi
23: 56690 Kigali
24: 1200745 Kivu
25: 25645 Rwanda

```
26: 1863662 Gisenyi
27: 1200745 Kivu
28: 1863662 Gisenyi
29: 25645 Rwanda
30: 1863662 Gisenyi
1: 25645 Rwanda
    [pem_pos=0, final_scores=0.48027748,
final_local_scores=0.41952628, pem=-0.42949,
similarity=18.944847]
2: 76762 Democratic_Republic_of_the_Congo
    [pem_pos=3, final_scores=0.35945135,
final_local_scores=0.32050797, pem=-2.0037596,
similarity=24.551067]
3: 25645 Rwanda
    [pem_pos=0, final_scores=0.6020685,
final_local_scores=0.55256444, pem=0.0, similarity=20.579338]
4: 56690 Kigali
    [pem_pos=0, final_scores=0.42931646,
final_local_scores=0.40953153, pem=0.0, similarity=16.310951]
5: 25645 Rwanda
    [pem_pos=0, final_scores=0.4621277,
final_local_scores=0.3997004, pem=-0.42949,
similarity=18.353205]
6: 25645 Rwanda
    [pem_pos=0, final_scores=0.33677053,
final_local_scores=0.26276678, pem=0.0, similarity=11.931198]
7: 59537 Hutu
    [pem_pos=0, final_scores=0.49773306,
final_local_scores=0.5030842, pem=0.0, similarity=19.102749]
8: 76762 Democratic_Republic_of_the_Congo
    [pem_pos=3, final_scores=0.34063327,
final_local_scores=0.29995212, pem=-2.0037596,
similarity=23.93764]
9: 25645 Rwanda
    [pem_pos=0, final_scores=0.5368846,
final_local_scores=0.48136097, pem=0.0, similarity=18.454485]
10: 18998750 Reuters
    [pem_pos=0, final_scores=0.31644934,
final_local_scores=0.39783323, pem=0.0, similarity=15.961849]
11: 56690 Kigali
    [pem_pos=0, final_scores=0.40599024,
final_local_scores=0.3840512, pem=0.0, similarity=15.550568]
12: 1210483 Goma
    [pem_pos=0, final_scores=0.41455454,
final_local_scores=0.40111107, pem=0.0, similarity=16.059668]
13: 59537 Hutu
    [pem_pos=0, final_scores=0.48871917,
final_local_scores=0.4932378, pem=0.0, similarity=18.808912]
14: 76762 Democratic_Republic_of_the_Congo
    [pem_pos=3, final_scores=0.32998973,
final_local_scores=0.28832564, pem=-2.0037596,
similarity=23.590681]
```

```
15: 67050 Tutsi
    [pem_pos=0, final_scores=0.4832844,
final_local_scores=0.4811442, pem=0.0, similarity=18.448015]
16: 25645 Rwanda
    [pem_pos=0, final_scores=0.60993546,
final_local_scores=0.56115794, pem=-0.42949,
similarity=23.17142]
17: 420869 Léon_Kengo
    [pem_pos=0, final_scores=0.7089301,
final_local_scores=0.82640857, pem=0.0, similarity=28.751392]
18: 25645 Rwanda
    [pem_pos=0, final_scores=0.5511365,
final_local_scores=0.496929, pem=-0.42949, similarity=21.2547]
19: 25645 Rwanda
    [pem_pos=0, final_scores=0.54057616,
final_local_scores=0.48539346, pem=-0.42949,
similarity=20.910456]
20: 25645 Rwanda
    [pem_pos=0, final_scores=0.47545367,
final_local_scores=0.414257, pem=0.0, similarity=16.451967]
21: 59537 Hutu
    [pem_pos=0, final_scores=0.4637205,
final_local_scores=0.46593058, pem=0.0, similarity=17.99401]
22: 67050 Tutsi
    [pem_pos=0, final_scores=0.42578787,
final_local_scores=0.418338, pem=0.0, similarity=16.573753]
23: 56690 Kigali
    [pem_pos=0, final_scores=0.3575068,
final_local_scores=0.33109045, pem=0.0, similarity=13.970113]
24: 1200745 Kivu
    [pem_pos=0, final_scores=0.3670482,
final_local_scores=0.38015127, pem=0.0, similarity=15.434186]
25: 25645 Rwanda
    [pem_pos=0, final_scores=0.54143065,
final_local_scores=0.4863268, pem=0.0, similarity=18.602674]
26: 1863662 Gisenyi
    [pem_pos=0, final_scores=0.5463834,
final_local_scores=0.57966757, pem=0.0, similarity=21.38815]
27: 1200745 Kivu
    [pem_pos=0, final_scores=0.33610898,
final_local_scores=0.3482648, pem=0.0, similarity=14.48263]
28: 1863662 Gisenyi
    [pem_pos=0, final_scores=0.54324925,
final_local_scores=0.576244, pem=0.0, similarity=21.285984]
29: 25645 Rwanda
    [pem_pos=0, final_scores=0.64074355,
final_local_scores=0.5948111, pem=0.0, similarity=21.840061]
30: 1863662 Gisenyi
    [pem_pos=0, final_scores=0.56280535,
final_local_scores=0.59760606, pem=0.0, similarity=21.92347]
1: 76762 Democratic_Republic_of_the_Congo
    [pem_pos=3, final_scores=0.0431834,
```

```
final_local_scores=-0.028781563, pem=-2.0037596,  
similarity=14.12757]  
2: 76762 Democratic_Republic_of_the_Congo  
    [pem_pos=3, final_scores=0.075390264,  
final_local_scores=0.010214418, pem=-2.0037596,  
similarity=15.291287]
```

Sample: b'286 Revered'

Revered skull of S. Africa king is [1 1 Scottish 1] woman's . [2 2 JOHANNESBURG 2] 1996-08-24 A limelight-loving [3 3 South African 3] chief was in disgrace on Saturday after a prized skull he brought home from [4 4 Scotland 4] was identified as belonging not to his sacred tribal ancestor , but to a middle-aged white woman . A forensic scientist who examined the supposed skull of 19th century King [5 5 Hintsa 5] , a chief of President [6 6 Nelson Mandela 6] 's [7 7 Xhosa 7] tribe killed in battle by the [8 8 British 8] , said it was in fact the cranium of a [9 9 European 9] woman . Chief Nicholas Gcaleka , dressed in animal skins and full tribal regalia , journeyed to a wintry [10 10 Scotland 10] in February on a hugely publicised quest to find [11 11 Hintsa 11] 's skull . The witchdoctor said ancestors had appeared to him in a dream and ordered him to return the head , said to have been carried off as a colonial trophy by the officer who shot and allegedly beheaded [12 12 Hintsa 12] after a battle in 1835 . But Gcaleka ran into trouble as soon as he returned to [13 13 South Africa 13] with a skull he found in a cottage in a lonely [14 14 Highland 14] forest near [15 15 Inverness 15] . He said the spirit of a hurricane had guided him there . Members of the [16 16 Xhosa 16] royal family , branding Gcaleka a charlatan , confiscated the head and sent it for tests to a forensic scientist , who examined the shape of the skull and the hole that he determined had not come , as supposed , from a bullet . " It can be stated beyond reasonable doubt that this skull is not that of the late king , " the scientist said in a statement .

```
1: 26994 Scotland  
2: 58893 Johannesburg  
3: 17416221 South_Africa  
4: 26994 Scotland  
5: 7161613 Hintsa_kakhawuta  
6: 21492751 Nelson_Mandela  
7: 19997199 Xhosa_people  
8: 31717 United_Kingdom  
9: 9239 Europe  
10: 26994 Scotland  
11: 7161613 Hintsa_kakhawuta  
12: 7161613 Hintsa_kakhawuta  
13: 17416221 South_Africa  
14: 28894 Scottish_Highlands  
15: 144968 Inverness  
16: 19997199 Xhosa_people
```

```
1: 26994 Scotland
   [pem_pos=0, final_scores=0.39579672,
final_local_scores=0.3653335, pem=-0.47951394,
similarity=17.599665]
2: 58893 Johannesburg
   [pem_pos=0, final_scores=0.46736372,
final_local_scores=0.4843738, pem=-0.09562059,
similarity=19.064394]
3: 17416221 South_Africa
   [pem_pos=0, final_scores=0.46444148,
final_local_scores=0.46991062, pem=-0.1389451,
similarity=18.868387]
4: 26994 Scotland
   [pem_pos=0, final_scores=0.4982713,
final_local_scores=0.47727144, pem=-0.22905664,
similarity=19.57809]
5: 7161613 Hintsa_kaKhawuta
   [pem_pos=0, final_scores=0.60689604,
final_local_scores=0.68397534, pem=0.0, similarity=24.500902]
6: 21492751 Nelson_Mandela
   [pem_pos=0, final_scores=0.43740273,
final_local_scores=0.4839201, pem=0.0, similarity=18.530853]
7: 19997199 Xhosa_people
   [pem_pos=2, final_scores=0.2808985,
final_local_scores=0.2796169, pem=-2.850098,
similarity=27.933315]
8: 31717 United_Kingdom
   [pem_pos=0, final_scores=0.32837194,
final_local_scores=0.32153016, pem=-0.48936236,
similarity=16.346043]
9: 9239 Europe
   [pem_pos=0, final_scores=0.40348667,
final_local_scores=0.4343233, pem=-0.25557294,
similarity=18.440632]
10: 26994 Scotland
    [pem_pos=0, final_scores=0.42243904,
final_local_scores=0.39443624, pem=-0.22905664,
similarity=17.106123]
11: 7161613 Hintsa_kaKhawuta
    [pem_pos=0, final_scores=0.56869924,
final_local_scores=0.64225113, pem=0.0, similarity=23.255768]
12: 7161613 Hintsa_kaKhawuta
    [pem_pos=0, final_scores=0.6164306,
final_local_scores=0.69439036, pem=0.0, similarity=24.811707]
13: 17416221 South_Africa
    [pem_pos=0, final_scores=0.49860495,
final_local_scores=0.507229, pem=-0.26251382,
similarity=20.654028]
14: 28894 Scottish_Highlands
    [pem_pos=1, final_scores=0.41095048,
final_local_scores=0.4202507, pem=-2.2532446,
similarity=28.884327]
```

```
15: 144968 Inverness
    [pem_pos=0, final_scores=0.6086333,
final_local_scores=0.62059045, pem=-0.19422773,
similarity=23.66561]
16: 19997199 Xhosa_people
    [pem_pos=2, final_scores=0.293904,
final_local_scores=0.29382342, pem=-2.850098,
similarity=28.357265]
```

Sample: b'287_Sudan'

[1 1 Sudan 1] arrests opposition sewing machine smugglers . [2 2 KHARTOUM 2] 1996-08-24 [3 3 Sudanese 3] police have arrested three people trying to smuggle sewing machines and army clothing to [4 4 Sudanese 4] opposition groups in [5 5 Eritrea 5] , an official newspaper reported on Saturday . The government-owned al-Ingaz al-Watani said the smugglers were caught in Banat in the eastern state of [6 6 Kassala 6] , on the border with [7 7 Eritrea 7] , and had confessed they were on their way to " the so-called alliance forces which have been undertaking subversive operations on the eastern border " . Authorities in [8 8 Kassala 8] said opposition forces based in [9 9 Eritrea 9] have been laying landmines and stealing vehicles and other goods to smuggle them across the border into [10 10 Eritrea 10] . [11 11 Sudan 11] accuses the [12 12 Eritrean 12] authorities of providing support to [13 13 Sudanese 13] opposition elements based in [14 14 Eritrea 14] . [15 15 Eritrea 15] cut diplomatic ties with [16 16 Sudan 16] in 1994 , accusing it of training rebels to make raids into [17 17 Eritrea 17] . The exiled [18 18 National Democratic Alliance 18] , a [19 19 Sudanese 19] umbrella opposition group , has its headquarters in the [20 20 Eritrean 20] capital [21 21 Asmara 21] . It uses the former [22 22 Sudanese 22] embassy .

1: 27421 Sudan

2: 17320 Khartoum

3: 27421 Sudan

4: 27421 Sudan

5: 17238590 Eritrea

6: 767657 Kassala_(state)

7: 17238590 Eritrea

8: 1504036 Kassala

9: 17238590 Eritrea

10: 17238590 Eritrea

11: 27421 Sudan

12: 17238590 Eritrea

13: 27421 Sudan

14: 17238590 Eritrea

15: 17238590 Eritrea

16: 27421 Sudan

17: 17238590 Eritrea

18: 2068118 National_Democratic_Alliance_(Sudan)

19: 27421 Sudan

20: 17238590 Eritrea

```
21: 56827 Asmara
22: 27421 Sudan
1: 27421 Sudan
    [pem_pos=0, final_scores=0.6015516,
final_local_scores=0.5426074, pem=-0.2848764,
similarity=21.831404]
2: 17320 Khartoum
    [pem_pos=0, final_scores=0.38242072,
final_local_scores=0.35850155, pem=-0.03966028,
similarity=15.003793]
3: 27421 Sudan
    [pem_pos=0, final_scores=0.5911105,
final_local_scores=0.53120214, pem=-0.19893567,
similarity=21.023687]
4: 27421 Sudan
    [pem_pos=0, final_scores=0.58703405,
final_local_scores=0.5267492, pem=-0.19893567,
similarity=20.890802]
5: 17238590 Eritrea
    [pem_pos=0, final_scores=0.5202735,
final_local_scores=0.45509595, pem=-0.3232545,
similarity=19.428593]
6: 767657 Kassala_(state)
    [pem_pos=1, final_scores=0.2697938,
final_local_scores=0.3324248, pem=-0.27964067,
similarity=15.530662]
7: 17238590 Eritrea
    [pem_pos=0, final_scores=0.55491143,
final_local_scores=0.49293256, pem=-0.3232545,
similarity=20.557713]
8: 1504036 Kassala
    [pem_pos=0, final_scores=0.26510185,
final_local_scores=0.2936738, pem=0.0, similarity=12.853525]
9: 17238590 Eritrea
    [pem_pos=0, final_scores=0.52996093,
final_local_scores=0.46567792, pem=-0.3232545,
similarity=19.744379]
10: 17238590 Eritrea
    [pem_pos=0, final_scores=0.5629,
final_local_scores=0.5016589, pem=-0.3232545,
similarity=20.818125]
11: 27421 Sudan
    [pem_pos=0, final_scores=0.58969355,
final_local_scores=0.5296543, pem=-0.2848764,
similarity=21.444857]
12: 17238590 Eritrea
    [pem_pos=0, final_scores=0.5459718,
final_local_scores=0.48316735, pem=0.0, similarity=18.50839]
13: 27421 Sudan
    [pem_pos=0, final_scores=0.5573427,
final_local_scores=0.49431592, pem=-0.19893567,
similarity=19.92293]
```

```
14: 17238590 Eritrea
    [pem_pos=0, final_scores=0.5612528,
final_local_scores=0.4998595, pem=-0.3232545,
similarity=20.764427]
15: 17238590 Eritrea
    [pem_pos=0, final_scores=0.4949634,
final_local_scores=0.42744845, pem=-0.3232545,
similarity=18.603537]
16: 27421 Sudan
    [pem_pos=0, final_scores=0.60678124,
final_local_scores=0.54832006, pem=-0.2848764,
similarity=22.001879]
17: 17238590 Eritrea
    [pem_pos=0, final_scores=0.57318795,
final_local_scores=0.5128969, pem=-0.3232545,
similarity=21.153488]
18: 2068118 National_Democratic_Alliance_(Sudan)
    [pem_pos=1, final_scores=0.33882803,
final_local_scores=0.44257897, pem=-1.8980871,
similarity=27.619246]
19: 27421 Sudan
    [pem_pos=0, final_scores=0.53969795,
final_local_scores=0.47504175, pem=-0.19893567,
similarity=19.34775]
20: 17238590 Eritrea
    [pem_pos=0, final_scores=0.50848466,
final_local_scores=0.44221842, pem=0.0, similarity=17.286392]
21: 56827 Asmara
    [pem_pos=0, final_scores=0.39335448,
final_local_scores=0.38074154, pem=0.0, similarity=15.451799]
22: 27421 Sudan
    [pem_pos=0, final_scores=0.5788931,
final_local_scores=0.51785654, pem=-0.19893567,
similarity=20.62543]
```

Sample: b'288_Albanian'

```
[1 1 Albanian 1] Socialists start landmark reform congress .
[2 2 TIRANA 2] 1996-08-24 [3 3 Albania 3] 's opposition [1 1 S
ocialist Party 1] began a two-day congress on Saturday to
discuss major jettisoning its links with almost half a
century of Stalinist dictatorship in the [4 4 Balkan 4]
country . " The congress will approve new concepts that will
turn the party into a Social-Democratic and electoral party ,
not a class and ideological one , " the Socialist [5 5 Zeri i
Popullit 5] daily said in an editorial . Jailed Socialist
leader [6 6 Fatos Nano 6] made the first call for change in
July , a month after the party 's chief opponents , the
conservative Democrats of President [7 7 Sali Berisha 7] ,
almost swept the board in a disputed general election . The
Socialists , reformed heirs to the communists , pulled out of
the poll saying it was a sham . Acting Socialist leader [8 8 S
erbet Pellumbi 8] has said he too will urge the party to
```

scrap the ideas of [9 9 Karl Marx 9] at the congress . The pro-reform stance of some of the party leadership initially caused a storm and triggered the resignation last month of the party 's Secretary-General Gramoz Ruci . More recently political commentators have reported a growing consensus , however , and a rift at the meeting looks increasingly unlikely .

```
1: 738 Albania
2: 31075 Tirana
3: 738 Albania
4: 4829 Balkans
5: 15865800 Zëri_i_Popullit
6: 393701 Fatos_Nano
7: 267929 Sali_Berisha
8: 28153727 Servet_Pëllumbi
9: 16743 Karl_Marx
1: 401469 Socialist_Party_of_Albania
    [winner: 417429 Socialist_Party_of_Serbia, pem_pos=19,
final_scores=0.10470782, final_local_scores=0.11790711,
pem=-4.5017786, similarity=32.08967|
     RECALL MISS: 401469 Socialist_Party_of_Albania]
1: 738 Albania
    [pem_pos=2, final_scores=0.5687157,
final_local_scores=0.5680508, pem=-1.5560378,
similarity=29.503458]
2: 31075 Tirana
    [pem_pos=0, final_scores=0.40024108,
final_local_scores=0.38324952, pem=0.0, similarity=15.526643]
3: 738 Albania
    [pem_pos=0, final_scores=0.55204105,
final_local_scores=0.5498363, pem=-0.41162592,
similarity=22.73641]
4: 4829 Balkans
    [pem_pos=0, final_scores=0.50622755,
final_local_scores=0.5687595, pem=-0.13899443,
similarity=21.818504]
5: 15865800 Zëri_i_Popullit
    [pem_pos=0, final_scores=0.65509945,
final_local_scores=0.7614854, pem=0.0, similarity=26.813955]
6: 393701 Fatos_Nano
    [pem_pos=0, final_scores=0.62911737,
final_local_scores=0.6659373, pem=0.0, similarity=23.96261]
7: 267929 Sali_Berisha
    [pem_pos=0, final_scores=0.8430374,
final_local_scores=0.89635354, pem=0.0, similarity=30.838692]
8: 28153727 Servet_Pëllumbi
    [pem_pos=0, final_scores=0.46168888,
final_local_scores=0.57901824, pem=0.0, similarity=21.368773]
9: 16743 Karl_Marx
    [pem_pos=0, final_scores=0.33262402,
final_local_scores=0.4377938, pem=0.0, similarity=17.154352]
1: 417429 Socialist_Party_of_Serbia
```

```
[pem_pos=19, final_scores=0.10470782,  
final_local_scores=0.11790711, pem=-4.5017786,  
similarity=32.08967]
```

Sample: b'289_Nicaraguan'

```
[1 1 Nicaraguan 1] president to go to [2 2 U.S. 2] for  
medical care . [3 3 MANAGUA 3] , [4 4 Nicaragua 4] 1996-08-23  
[5 5 Nicaraguan 5] President [6 6 Violeta Chamorro 6] was due  
to fly to the [7 7 United States 7] on Saturday for a medical  
check-up to determine if surgery was needed on the lower part  
of her spinal column , the government said on Friday . [8 8 Ch  
amorro 8] has complained of lower back pain since her trip to  
[9 9 Taiwan 9] in May , when the pain forced her to go to  
Taipei University Hospital for an examination . [10 10 Chamor  
ro 10] , 66 , suffers from osteoporosis , a disease that  
weakens the bones , and has repeatedly flown to [11 11 Washin  
gton 11] for treatment by her longtime doctor , Sam Wilson .
```

1: 21362 Nicaragua

2: 3434750 United_States

3: 57042 Managua

4: 21362 Nicaragua

5: 21362 Nicaragua

6: 264238 Violeta_Chamorro

7: 3434750 United_States

8: 264238 Violeta_Chamorro

9: 25734 Taiwan

10: 264238 Violeta_Chamorro

11: 108956 Washington,_D.C.

1: 21362 Nicaragua

```
[pem_pos=0, final_scores=0.512989,  
final_local_scores=0.48447388, pem=0.0, similarity=18.54738]
```

2: 3434750 United_States

```
[pem_pos=0, final_scores=0.4237095,  
final_local_scores=0.4337706, pem=-0.14075686,  
similarity=17.799751]
```

3: 57042 Managua

```
[pem_pos=0, final_scores=0.41274005,  
final_local_scores=0.39977884, pem=0.0, similarity=16.01991]
```

4: 21362 Nicaragua

```
[pem_pos=0, final_scores=0.5278375,  
final_local_scores=0.5006936, pem=-0.24729352,  
similarity=20.37623]
```

5: 21362 Nicaragua

```
[pem_pos=0, final_scores=0.481866,  
final_local_scores=0.4504767, pem=0.0, similarity=17.532837]
```

6: 264238 Violeta_Chamorro

```
[pem_pos=0, final_scores=0.8439686,  
final_local_scores=0.86251616, pem=0.0, similarity=29.828917]
```

7: 3434750 United_States

```
[pem_pos=0, final_scores=0.47261614,  
final_local_scores=0.48719364, pem=-0.073457584,  
similarity=19.028017]
```

```
8: 264238 Violeta_Chamorro
    [pem_pos=0, final_scores=0.8463362,
final_local_scores=0.879314, pem=0.0, similarity=30.330198]
9: 25734 Taiwan
    [pem_pos=0, final_scores=0.342588,
final_local_scores=0.37239307, pem=-0.077886656,
similarity=15.626225]
10: 264238 Violeta_Chamorro
    [pem_pos=0, final_scores=0.7931742,
final_local_scores=0.8212425, pem=0.0, similarity=28.597227]
11: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.4612838,
final_local_scores=0.47868198, pem=-1.7496481,
similarity=27.889397]
```

Sample: b'290_Nepal'

[1 1 Nepal 1] wo n't help split [2 2 Tibet 2] , king tells [3 3 China 3] . [4 4 BEIJING 4] 1996-08-24 King [5 5 Birendra 5] of [6 6 Nepal 6] has told [7 7 China 7] his nation will not become the tool of people who want [8 8 Tibetan 8] independence from [9 9 Beijing 9] , the official [10 10 China Daily 10] newspaper said on Saturday . King [11 11 Birendra 11] , in [12 12 Tibet 12] at the start of a one-week unofficial visit to [13 13 China 13] , said the [14 14 Nepalese 14] government had " maintained a sharp vigilance against such intentions " , the newspaper said . [15 15 Nepal 15] shares a long mountain border with the restive [16 16 Himalayan 16] region , where opposition to [17 17 Beijing 17] 's four-decade rule is widespread . [18 18 Chinese 18] official media has often accused foreign forces , notably the [19 19 United States 19] , of seeking to support [20 20 Tibetan 20] independence activists . King [21 21 Birendra 21] told Gyaicain Norbu , chairman of the [22 22 Tibetan 22] government , that [23 23 Nepal 23] would not " become a tool for others to split [24 24 Tibet 24] " , the newspaper said . Gyaicain told the royal visitor increased cooperation between [25 25 Nepal 25] and [26 26 Tibet 26] was possible in the fields of trade , tourism , communications and sports , it said . It gave no details .

```
1: 171166 Nepal
2: 31516 Tibet
3: 5405 China
4: 18603746 Beijing
5: 313169 Birendra_of_Nepal
6: 171166 Nepal
7: 5405 China
8: 31516 Tibet
9: 18603746 Beijing
10: 737850 China_Daily
11: 313169 Birendra_of_Nepal
12: 31516 Tibet
13: 5405 China
```

14: 171166 Nepal
15: 171166 Nepal
16: 37966 Himalayas
17: 18603746 Beijing
18: 5405 China
19: 3434750 United_States
20: 31516 Tibet
21: 313169 Birendra_of_Nepal
22: 31516 Tibet
23: 171166 Nepal
24: 31516 Tibet
25: 171166 Nepal
26: 31516 Tibet
1: 171166 Nepal
 [pem_pos=1, final_scores=0.55408186,
final_local_scores=0.50812143, pem=-0.22887385,
similarity=20.49772]
2: 31516 Tibet
 [pem_pos=0, final_scores=0.5011116,
final_local_scores=0.4497071, pem=-0.107068084,
similarity=18.092121]
3: 5405 China
 [pem_pos=0, final_scores=0.5200998,
final_local_scores=0.48084557, pem=-0.15669827,
similarity=19.291254]
4: 18603746 Beijing
 [pem_pos=0, final_scores=0.50420505,
final_local_scores=0.5202946, pem=-0.042161036,
similarity=19.845621]
5: 313169 Birendra_of_Nepal
 [pem_pos=0, final_scores=0.5574408,
final_local_scores=0.6193779, pem=0.0, similarity=22.573185]
6: 171166 Nepal
 [pem_pos=0, final_scores=0.60233307,
final_local_scores=0.56082857, pem=-0.22887385,
similarity=22.070608]
7: 5405 China
 [pem_pos=0, final_scores=0.524009,
final_local_scores=0.4851157, pem=-0.15669827,
similarity=19.418684]
8: 31516 Tibet
 [pem_pos=1, final_scores=0.48595148,
final_local_scores=0.43314695, pem=-1.2825032,
similarity=23.990135]
9: 18603746 Beijing
 [pem_pos=0, final_scores=0.4142143,
final_local_scores=0.42199343, pem=-0.042161036,
similarity=16.912117]
10: 737850 China_Daily
 [pem_pos=0, final_scores=0.53846043,
final_local_scores=0.6153963, pem=0.0, similarity=22.454367]
11: 313169 Birendra_of_Nepal

[pem_pos=0, final_scores=0.4754082,
final_local_scores=0.5297698, pem=0.0, similarity=19.8991]
12: 31516 Tibet
[pem_pos=0, final_scores=0.48908603,
final_local_scores=0.43657094, pem=-0.107068084,
similarity=17.700113]
13: 5405 China
[pem_pos=0, final_scores=0.50352454,
final_local_scores=0.4627396, pem=-0.15669827,
similarity=18.750935]
14: 171166 Nepal
[pem_pos=0, final_scores=0.6383777,
final_local_scores=0.6002019, pem=-0.16643417,
similarity=22.90603]
15: 171166 Nepal
[pem_pos=1, final_scores=0.64870685,
final_local_scores=0.6114849, pem=-0.22887385,
similarity=23.582293]
16: 37966 Himalayas
[pem_pos=0, final_scores=0.4122892,
final_local_scores=0.39728367, pem=0.0, similarity=15.94545]
17: 18603746 Beijing
[pem_pos=0, final_scores=0.42986053,
final_local_scores=0.43908453, pem=-0.042161036,
similarity=17.42215]
18: 5405 China
[pem_pos=1, final_scores=0.6383676,
final_local_scores=0.6100351, pem=-1.4196125,
similarity=30.014452]
19: 3434750 United_States
[pem_pos=0, final_scores=0.48993272,
final_local_scores=0.5150883, pem=-0.073457584,
similarity=19.86045]
20: 31516 Tibet
[pem_pos=1, final_scores=0.5259934,
final_local_scores=0.47688663, pem=-1.2825032,
similarity=25.295414]
21: 313169 Birendra_of_Nepal
[pem_pos=0, final_scores=0.5962502,
final_local_scores=0.6617713, pem=0.0, similarity=23.83829]
22: 31516 Tibet
[pem_pos=1, final_scores=0.49851698,
final_local_scores=0.44687283, pem=-1.2825032,
similarity=24.399742]
23: 171166 Nepal
[pem_pos=1, final_scores=0.5952588,
final_local_scores=0.553101, pem=-0.22887385,
similarity=21.840002]
24: 31516 Tibet
[pem_pos=0, final_scores=0.48478287,
final_local_scores=0.4318704, pem=-0.107068084,
similarity=17.55984]

```
25: 171166 Nepal
    [pem_pos=1, final_scores=0.6606726,
final_local_scores=0.62455565, pem=-0.22887385,
similarity=23.972351]
26: 31516 Tibet
    [pem_pos=0, final_scores=0.43965238,
final_local_scores=0.38257217, pem=-0.107068084,
similarity=16.088682]

Sample: b'291_Iran'
[1 1 Iran 1] hangs two men for drug trafficking . [2 2 TEHRAN
2] 1996-08-24 [3 3 Iran 3] has hanged two drug traffickers in
the southern city of [4 4 Shiraz 4] , the evening newspaper [5
5 Resalat 5] reported on Saturday . The two [6 6 Iranian 6]
men were arrested in July with 419 kilograms ( 924 lbs ) of
opium after they opened fire on police and killed a
pedestrain and wounded four , the newspaper quoted a police
commander as saying . [7 7 Resalat 7] said the executions
were ordered by the Islamic Revolutionary Court . It did not
say when they took place . One of the men , who killed the
pedestrian , was hanged at the site of the crime and the
other was executed in Adel prison in [8 8 Shiraz 8] , the
newspaper said . Possession of 30 grammes ( just over an
ounce ) of heroin or five kg ( 11 lb ) of opium is punishable
by death in [9 9 Iran 9] . More than 1,000 people have been
executed in drug-related cases since the law took effect in
1989 . [10 10 Iran 10] has an estimated one million drug
addicts and is a key transit route for drugs , mostly opium ,
smuggled to [11 11 Europe 11] through [12 12 Afghanistan 12]
and [13 13 Pakistan 13] -- the so called " Golden Crescent . "
1: 14653 Iran
2: 57654 Tehran
3: 14653 Iran
4: 39622 Shiraz
5: 24976888 Resalat_(newspaper)
6: 14653 Iran
7: 24976888 Resalat_(newspaper)
8: 39622 Shiraz
9: 14653 Iran
10: 14653 Iran
11: 9239 Europe
12: 737 Afghanistan
13: 23235 Pakistan
1: 14653 Iran
    [pem_pos=0, final_scores=0.6034613,
final_local_scores=0.5813007, pem=-0.13856095,
similarity=22.190401]
2: 57654 Tehran
    [pem_pos=0, final_scores=0.42002344,
final_local_scores=0.41006655, pem=0.0, similarity=16.326916]
3: 14653 Iran
    [pem_pos=0, final_scores=0.7097575,
```

```
final_local_scores=0.69741315, pem=-0.13856095,
similarity=25.655428]
4: 39622 Shiraz
    [pem_pos=0, final_scores=0.35505944,
final_local_scores=0.34000075, pem=-0.105005756,
similarity=14.807051]
5: 24976888 Resalat_(newspaper)
    [pem_pos=1, final_scores=0.43720955,
final_local_scores=0.55479985, pem=-0.6523252,
similarity=24.193493]
6: 14653 Iran
    [pem_pos=0, final_scores=0.5924595,
final_local_scores=0.5692829, pem=-0.57066226,
similarity=24.181599]
7: 24976888 Resalat_(newspaper)
    [pem_pos=1, final_scores=0.4149955,
final_local_scores=0.53053445, pem=-0.6523252,
similarity=23.469366]
8: 39622 Shiraz
    [pem_pos=0, final_scores=0.34977257,
final_local_scores=0.33422565, pem=-0.105005756,
similarity=14.634711]
9: 14653 Iran
    [pem_pos=0, final_scores=0.4933092,
final_local_scores=0.4609763, pem=-0.13856095,
similarity=18.599682]
10: 14653 Iran
    [pem_pos=0, final_scores=0.53238225,
final_local_scores=0.50365764, pem=-0.13856095,
similarity=19.873377]
11: 9239 Europe
    [pem_pos=0, final_scores=0.54123443,
final_local_scores=0.6184403, pem=-0.086338006,
similarity=23.014725]
12: 737 Afghanistan
    [pem_pos=0, final_scores=0.57246196,
final_local_scores=0.57715726, pem=-0.22286008,
similarity=22.525185]
13: 23235 Pakistan
    [pem_pos=0, final_scores=0.49775332,
final_local_scores=0.49820602, pem=-0.21944217,
similarity=20.150534]
```

Sample: b'292_Main'

Main [1 1 Tunisian 1] opposition party ousted from HQ . [2 2 TUNIS 2] 1996-08-24 [3 3 Tunisia 3] 's main opposition party on Saturday announced that it had been ousted from its headquarters building by a court decision for failing to pay the rent . Mohamed Ali Khalfallah , spokesman for the [4 4 Movement of Socialist Democrats 4] (MDS) said that a bailiff who was accompanied by policemen , on Saturday ordered the party to leave the building . " We were not allowed a delay

to enable us to transfer the movement 's goods and documents , " Khalfallah added in a statement . The building is state property . The MDS was represented in court and admitted owing money for rent but did not give details . MDS this year lost its president and vice-president , both of whom were tried and given jail sentences . MDS president Mohamed Moada was sentenced last February to 11 years in jail on charges of having secret contacts with [5 5 Libyan 5] agents and receiving money from Tripoli . Vice-president Khemais Chamhari last July was sentenced to five years in prison on a charge of disclosing secrets of judicial proceedings in Moada 's affair . To replace Moada , the MDS after the trial named Khalfallah as " coordinator " but Ismail Boulahya , the last of the MDS founding members still politically active , claimed the title of president , causing a new split within the movement . MDS was founded in 1978 by a group led by Ahmed Mestiri , who withdrew from politics in 1992 . Succeeding him as head of the movement , Moada , an [6 6 Arab 6] nationalist , ousted liberals led by MDS secretary-general Mustapha Ben Jaafar in 1993 .

1: 30188 Tunisia
2: 57659 Tunis
3: 30188 Tunisia
4: 403110 Movement_of_Socialist_Democrats
5: 17633 Libya
6: 159433 Arab_world
1: 30188 Tunisia
[pem_pos=0, final_scores=0.5464303,
final_local_scores=0.4955048, pem=-0.13068858,
similarity=19.587269]
2: 57659 Tunis
[pem_pos=0, final_scores=0.4193673,
final_local_scores=0.39541805, pem=-0.032719024,
similarity=16.067707]
3: 30188 Tunisia
[pem_pos=0, final_scores=0.5046238,
final_local_scores=0.44983757, pem=-0.39291763,
similarity=19.65051]
4: 403110 Movement_of_Socialist_Democrats
[pem_pos=0, final_scores=0.43886626,
final_local_scores=0.56953603, pem=0.0, similarity=21.085804]
5: 17633 Libya
[pem_pos=0, final_scores=0.65980786,
final_local_scores=0.64593273, pem=-0.27370736,
similarity=24.854097]
6: 159433 Arab_world
[pem_pos=14, final_scores=0.5705305,
final_local_scores=0.58828175, pem=-3.8052368,
similarity=42.338684]

Sample: b'293_Kurdish'

[1 1 Kurdish 1] group says two killed in [2 2 Iraqi 2]

shelling . [3 3 NICOSIA 3] 1996-08-24 An [4 4 Iraqi Kurdish 4] guerrilla group on Saturday accused [5 5 Iraqi 5] government forces of killing two civilians in shelling in northern [6 6 Iraq 6] , the [7 7 Iranian 7] news agency [8 8 IRNA 8] reported . [9 9 IRNA 9] said it was monitoring a report from a radio station affiliated to the [10 10 Patriotic Union of Kurdistan 10] ([11 11 PUK 11]) . " [12 12 Iraqi 12] army heavily shelled the Kanie Karzhala camp , west of [13 13 Arbil 13] , on Friday ... Two civilians were killed in the [14 14 Iraqi 14] bombing , " [15 15 IRNA 15] quoted the radio report as saying . The PUK-run radio on Friday said [16 16 Iraqi 16] heavy artillery was pounding its positions in Kurdish-controlled northern [17 17 Iraq 17] but it gave no details of casualties . There was no independent confirmation of the reports . The rival [18 18 Kurdistan Democratic Party 18] ([19 19 KDP 19]) , which accuses [20 20 Iran 20] of supporting the [21 21 PUK 21] , said on Thursday that its forces had halted an Iranian-backed attack by thousands of [22 22 PUK 22] fighters . The [23 23 United States 23] said in [24 24 Washington 24] on Friday that it had brokered a ceasefire to end six days of fighting between the two main [25 25 Kurdish 25] factions and persuaded them to attend U.S.-mediated peace talks next month . The clashes , shattering a ceasefire negotiated last year by 1 Washington , had threatened a U.S.-led peace plan to unite the [26 26 Kurdish 26] region against [27 27 Iraqi 27] President [28 28 Saddam Hussein 28] . [29 29 U.S. 29] , [30 30 British 30] and [31 31 French 31] planes have been patrolling the skies of northern [32 32 Iraq 32] since shortly after the 1991 [33 33 Gulf War 33] to shield [34 34 Iraq 34] 's [35 35 Kurds 35] from any attack by [36 36 Iraqi 36] troops .

1: 80777 Kurdistan

2: 7515928 Iraq

3: 57064 Nicosia

4: 679693 Iraqi_Kurdistan

5: 7515928 Iraq

6: 7515928 Iraq

7: 14653 Iran

8: 841314 Islamic_Republic_News_Agency

9: 841314 Islamic_Republic_News_Agency

10: 70789 Patriotic_Union_of_Kurdistan

11: 70789 Patriotic_Union_of_Kurdistan

12: 7515928 Iraq

13: 24534207 Erbil

14: 7515928 Iraq

15: 841314 Islamic_Republic_News_Agency

16: 7515928 Iraq

17: 7515928 Iraq

18: 25659771 Kurdistan_Democratic_Party

19: 25659771 Kurdistan_Democratic_Party

20: 14653 Iran

21: 70789 Patriotic_Union_of_Kurdistan

22: 70789 Patriotic_Union_of_Kurdistan
23: 3434750 United_States
24: 108956 Washington,_D.C.
25: 80777 Kurdistan
26: 80777 Kurdistan
27: 7515928 Iraq
28: 29490 Saddam_Hussein
29: 3434750 United_States
30: 31717 United_Kingdom
31: 5843419 France
32: 7515928 Iraq
33: 182000 Gulf_War
34: 7515928 Iraq
35: 80777 Kurdistan
36: 7515928 Iraq

1: 3434750 United_States
[winner: 108956 Washington,_D.C., pem_pos=0,
final_scores=-0.16014898, final_local_scores=-0.18352148,
pem=-1.7496481, similarity=8.127924]
RECALL MISS: 3434750 United_States]

1: 80777 Kurdistan
[pem_pos=9, final_scores=0.31006986,
final_local_scores=0.25397444, pem=-2.8001814,
similarity=26.896639]

2: 7515928 Iraq
[pem_pos=0, final_scores=0.5158389,
final_local_scores=0.4534986, pem=-0.48469657,
similarity=20.258871]

3: 57064 Nicosia
[pem_pos=0, final_scores=0.424501,
final_local_scores=0.45667785, pem=0.0, similarity=17.717892]

4: 679693 Iraqi_Kurdistan
[pem_pos=0, final_scores=0.46474642,
final_local_scores=0.4268688, pem=0.0, similarity=16.82833]

5: 7515928 Iraq
[pem_pos=0, final_scores=0.60106885,
final_local_scores=0.5465993, pem=-0.48469657,
similarity=23.037182]

6: 7515928 Iraq
[pem_pos=0, final_scores=0.6346914,
final_local_scores=0.5833269, pem=-0.25379938,
similarity=22.877552]

7: 14653 Iran
[pem_pos=0, final_scores=0.5233963,
final_local_scores=0.57277375, pem=-0.57066226,
similarity=24.285774]

8: 841314 Islamic_Republic_News_Agency
[pem_pos=0, final_scores=0.66817355,
final_local_scores=0.72121674, pem=0.0, similarity=25.612259]

9: 841314 Islamic_Republic_News_Agency
[pem_pos=0, final_scores=0.42899632,
final_local_scores=0.45995194, pem=0.0, similarity=17.815596]

10: 70789 Patriotic_Union_of_Kurdistan
[pem_pos=0, final_scores=0.6633238,
final_local_scores=0.6897206, pem=0.0, similarity=24.67235]
11: 70789 Patriotic_Union_of_Kurdistan
[pem_pos=0, final_scores=0.57804245,
final_local_scores=0.59656364, pem=0.0, similarity=21.892363]
12: 7515928 Iraq
[pem_pos=0, final_scores=0.5241736,
final_local_scores=0.46260297, pem=-0.48469657,
similarity=20.530563]
13: 24534207 Erbil
[pem_pos=0, final_scores=0.37290114,
final_local_scores=0.3538244, pem=0.0, similarity=14.648539]
14: 7515928 Iraq
[pem_pos=0, final_scores=0.5299589,
final_local_scores=0.4689225, pem=-0.48469657,
similarity=20.71915]
15: 841314 Islamic_Republic_News_Agency
[pem_pos=0, final_scores=0.60148424,
final_local_scores=0.6483688, pem=0.0, similarity=23.43833]
16: 7515928 Iraq
[pem_pos=0, final_scores=0.54279965,
final_local_scores=0.48294908, pem=-0.48469657,
similarity=21.137732]
17: 7515928 Iraq
[pem_pos=0, final_scores=0.61860746,
final_local_scores=0.5657576, pem=-0.25379938,
similarity=22.353249]
18: 25659771 Kurdistan_Democratic_Party
[pem_pos=0, final_scores=0.6336491,
final_local_scores=0.64763325, pem=0.0, similarity=23.416382]
19: 25659771 Kurdistan_Democratic_Party
[pem_pos=0, final_scores=0.62850213,
final_local_scores=0.6420109, pem=0.0, similarity=23.2486]
20: 14653 Iran
[pem_pos=0, final_scores=0.63231784,
final_local_scores=0.691754, pem=-0.13856095,
similarity=25.486547]
21: 70789 Patriotic_Union_of_Kurdistan
[pem_pos=0, final_scores=0.53622735,
final_local_scores=0.550887, pem=0.0, similarity=20.52928]
22: 70789 Patriotic_Union_of_Kurdistan
[pem_pos=0, final_scores=0.49482983,
final_local_scores=0.50566643, pem=0.0, similarity=19.179808]
23: 3434750 United_States
[pem_pos=0, final_scores=0.45019704,
final_local_scores=0.47025895, pem=-0.073457584,
similarity=18.522652]
24: 108956 Washington,_D.C.
[pem_pos=0, final_scores=0.3402714,
final_local_scores=0.37009388, pem=-1.7496481,
similarity=24.64891]

```
25: 80777 Kurdistan
    [pem_pos=9, final_scores=0.2661209,
final_local_scores=0.20596689, pem=-2.8001814,
similarity=25.463997]
26: 80777 Kurdistan
    [pem_pos=9, final_scores=0.39639795,
final_local_scores=0.34827477, pem=-2.8001814,
similarity=29.710749]
27: 7515928 Iraq
    [pem_pos=0, final_scores=0.53517896,
final_local_scores=0.47462463, pem=-0.48469657,
similarity=20.889315]
28: 29490 Saddam_Hussein
    [pem_pos=0, final_scores=0.46962637,
final_local_scores=0.45074576, pem=0.0, similarity=17.540865]
29: 3434750 United_States
    [pem_pos=0, final_scores=0.46165472,
final_local_scores=0.4827748, pem=-0.14075686,
similarity=19.262133]
30: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.59134465,
final_local_scores=0.61931396, pem=-0.48936236,
similarity=25.232506]
31: 5843419 France
    [pem_pos=0, final_scores=0.5340846,
final_local_scores=0.57467353, pem=-0.80223274,
similarity=25.601784]
32: 7515928 Iraq
    [pem_pos=0, final_scores=0.545175,
final_local_scores=0.48554385, pem=-0.25379938,
similarity=19.95951]
33: 182000 Gulf_War
    [pem_pos=0, final_scores=0.3495019,
final_local_scores=0.33541232, pem=0.0, similarity=14.099086]
34: 7515928 Iraq
    [pem_pos=0, final_scores=0.55713207,
final_local_scores=0.49860507, pem=-0.25379938,
similarity=20.349283]
35: 80777 Kurdistan
    [pem_pos=7, final_scores=0.30480617,
final_local_scores=0.24822465, pem=-2.482989,
similarity=25.000113]
36: 7515928 Iraq
    [pem_pos=0, final_scores=0.58148986,
final_local_scores=0.5252122, pem=-0.48469657,
similarity=22.398949]
```

Sample: b'294_Iran'

```
[1 1 Iran 1] accuses [2 2 Iraq 2] of ceasefire violations . [
3 3 NICOSIA 3] 1996-08-24 [4 4 Iran 4] has accused [5 5 Iraq 5
] of violating the ceasefire ending their 1980-88 war some 32
times between the end of March and May 31 this year , the [6 6
```

Iranian 6] news agency [7 7 IRNA 7] reported on Saturday . [8 8 Iran 8] 's deputy representative to the [9 9 United Nations 9] , Majid Takht Ravanchi , made the allegations in a letter to [10 10 U.N. 10] Secretary-General [11 11 Boutros Boutros-Ghali 11] on Friday , the agency said . " The Islamic Republic of Iran has reported some 32 new cases of ceasefire violations by the [12 12 Iraqi 12] regime between March 31 and May 31 , 1996 , " it reported from [13 13 New York 13] . It said violations included constructing observation posts , installing mortars and anti-aircraft cannons , setting up tents , penetrating [14 14 Iranian 14] territory , and firing rifle grenades towards [15 15 Iranian 15] territory . The eight-year war between the two countries ended with a U.N.-sponsored ceasefire .

1: 14653 Iran
2: 7515928 Iraq
3: 57064 Nicosia
4: 14653 Iran
5: 7515928 Iraq
6: 14653 Iran
7: 841314 Islamic_Republic_News_Agency
8: 14653 Iran
9: 31769 United_Nations
10: 31769 United_Nations
11: 211193 Boutros_Boutros-Ghali
12: 7515928 Iraq
13: 645042 New_York_City
14: 14653 Iran
15: 14653 Iran
1: 14653 Iran
[pem_pos=0, final_scores=0.59074885,
final_local_scores=0.5782531, pem=-0.13856095,
similarity=22.099455]
2: 7515928 Iraq
[pem_pos=0, final_scores=0.60967,
final_local_scores=0.5888502, pem=-0.25379938,
similarity=23.042377]
3: 57064 Nicosia
[pem_pos=0, final_scores=0.3780219,
final_local_scores=0.42634642, pem=0.0, similarity=16.81274]
4: 14653 Iran
[pem_pos=0, final_scores=0.6741615,
final_local_scores=0.6693688, pem=-0.13856095,
similarity=24.81853]
5: 7515928 Iraq
[pem_pos=0, final_scores=0.57817656,
final_local_scores=0.5544485, pem=-0.25379938,
similarity=22.015762]
6: 14653 Iran
[pem_pos=0, final_scores=0.57532376,
final_local_scores=0.5614036, pem=-0.57066226,
similarity=23.946465]

```
7: 841314 Islamic_Republic_News_Agency
    [pem_pos=0, final_scores=0.6578413,
final_local_scores=0.7390186, pem=0.0, similarity=26.143503]
8: 14653 Iran
    [pem_pos=0, final_scores=0.60871106,
final_local_scores=0.5978741, pem=-0.13856095,
similarity=22.684984]
9: 31769 United_Nations
    [pem_pos=0, final_scores=0.41078234,
final_local_scores=0.42546958, pem=0.0, similarity=16.786573]
10: 31769 United_Nations
    [pem_pos=0, final_scores=0.5540201,
final_local_scores=0.58193505, pem=0.0, similarity=21.455816]
11: 211193 Boutros_Boutros-Ghali
    [pem_pos=0, final_scores=0.5768481,
final_local_scores=0.67430735, pem=0.0, similarity=24.212389]
12: 7515928 Iraq
    [pem_pos=0, final_scores=0.5685808,
final_local_scores=0.5439666, pem=-0.48469657,
similarity=22.958616]
13: 645042 New_York_City
    [pem_pos=1, final_scores=0.35093045,
final_local_scores=0.42735702, pem=-1.7906687,
similarity=26.580833]
14: 14653 Iran
    [pem_pos=0, final_scores=0.5674582,
final_local_scores=0.5528117, pem=-0.57066226,
similarity=23.690065]
15: 14653 Iran
    [pem_pos=0, final_scores=0.5402288,
final_local_scores=0.52306765, pem=-0.57066226,
similarity=22.802444]
```

Sample: b'295_PRESS'

PRESS DIGEST - [1 1 Iraq 1] - Aug 24 . [2 2 BAGHDAD 2]
1996-08-24 These are some of the leading stories in the
official [3 3 Iraqi 3] press on Saturday . [4 4 Reuters 4]
has not verified these stories and does not vouch for their
accuracy . JUMHOURIYA - [5 5 Istanbul 5] chamber of commerce
urges [6 6 Ankara 6] to resume trade with [7 7 Iraq 7] . -
Offers from [8 8 Arab 8] and foreign companies to supply [9 9
Iraq 9] with goods . - Four ships unload tonnes of Iraq-bound
sugar at [10 10 Jordan 10] 's [11 11 Aqaba 11] . - Editorial
blames [12 12 U.S. 12] for latest flare-up of fighting
between [13 13 Kurdish 13] rebels in northern [14 14 Iraq 14]
. - Black market booms in the shadow of state-run
supermarkets . - Parliament completes draft law on protection
of river waters in [15 15 Iraq 15] . QADISSIYA - [16 16 Iraq 1
6] denounces violation of airspace by [17 17 U.S. 17]
warplanes . [18 18 IRAQ 18] - Editorial lambasts [19 19 Jalal
Talabani 19] , leader of a [20 20 Kurdish 20] rebel faction
in the north , for liaising with [21 21 Iran 21] in its fight

against rivals . [22 22 BABEL 22] - Blaming [23 23 Iraq 23] for riots in [24 24 Jordan 24] is a dirty game .

1: 7515928 Iraq
2: 4492 Baghdad
3: 7515928 Iraq
4: 18998750 Reuters
5: 3391396 Istanbul
6: 802 Ankara
7: 7515928 Iraq
8: 159433 Arab_world
9: 7515928 Iraq
10: 7515964 Jordan
11: 328399 Aqaba
12: 3434750 United_States
13: 80777 Kurdistan
14: 7515928 Iraq
15: 7515928 Iraq
16: 7515928 Iraq
17: 3434750 United_States
18: 7515928 Iraq
19: 440244 Jalal_Talabani
20: 80777 Kurdistan
21: 14653 Iran
22: 198957 Babel_(newspaper)
23: 7515928 Iraq
24: 7515964 Jordan
1: 7515928 Iraq
[pem_pos=0, final_scores=0.6014063,
final_local_scores=0.5404135, pem=-0.25379938,
similarity=21.59693]
2: 4492 Baghdad
[pem_pos=0, final_scores=0.43669426,
final_local_scores=0.38816404, pem=0.0, similarity=15.673303]
3: 7515928 Iraq
[pem_pos=0, final_scores=0.5536735,
final_local_scores=0.48827273, pem=-0.48469657,
similarity=21.2966]
4: 18998750 Reuters
[pem_pos=0, final_scores=0.33187056,
final_local_scores=0.3546807, pem=0.0, similarity=14.674092]
5: 3391396 Istanbul
[pem_pos=0, final_scores=0.4688849,
final_local_scores=0.46862435, pem=-0.025284406,
similarity=18.211899]
6: 802 Ankara
[pem_pos=0, final_scores=0.36487836,
final_local_scores=0.36157578, pem=0.0, similarity=14.879855]
7: 7515928 Iraq
[pem_pos=0, final_scores=0.68814564,
final_local_scores=0.6351631, pem=-0.25379938,
similarity=24.424448]
8: 159433 Arab_world

```
[pem_pos=14, final_scores=0.47485024,
final_local_scores=0.46416867, pem=-3.8052368,
similarity=38.634903]
9: 7515928 Iraq
    [pem_pos=0, final_scores=0.64650625,
final_local_scores=0.5896784, pem=-0.25379938,
similarity=23.067093]
10: 7515964 Jordan
    [pem_pos=0, final_scores=0.56300956,
final_local_scores=0.5161681, pem=-0.53647494,
similarity=22.410633]
11: 328399 Aqaba
    [pem_pos=0, final_scores=0.23866843,
final_local_scores=0.245895, pem=0.0, similarity=11.427711]
12: 3434750 United_States
    [pem_pos=0, final_scores=0.4561984,
final_local_scores=0.4690116, pem=-0.14075686,
similarity=18.851414]
13: 80777 Kurdistan
    [pem_pos=9, final_scores=0.29502857,
final_local_scores=0.240134, pem=-2.8001814,
similarity=26.483612]
14: 7515928 Iraq
    [pem_pos=0, final_scores=0.650895,
final_local_scores=0.59447247, pem=-0.25379938,
similarity=23.210157]
15: 7515928 Iraq
    [pem_pos=0, final_scores=0.595046,
final_local_scores=0.53346586, pem=-0.25379938,
similarity=21.389599]
16: 7515928 Iraq
    [pem_pos=0, final_scores=0.6060068,
final_local_scores=0.5454389, pem=-0.25379938,
similarity=21.746899]
17: 3434750 United_States
    [pem_pos=0, final_scores=0.45752525,
final_local_scores=0.47046095, pem=-0.14075686,
similarity=18.894665]
18: 7515928 Iraq
    [pem_pos=0, final_scores=0.537682,
final_local_scores=0.47080445, pem=-0.25379938,
similarity=19.519657]
19: 440244 Jalal_Talabani
    [pem_pos=0, final_scores=0.46484882,
final_local_scores=0.49636465, pem=0.0, similarity=18.902224]
20: 80777 Kurdistan
    [pem_pos=9, final_scores=0.30581164,
final_local_scores=0.25191283, pem=-2.8001814,
similarity=26.835115]
21: 14653 Iran
    [pem_pos=0, final_scores=0.7004654,
final_local_scores=0.7686886, pem=-0.13856095,
```

```
similarity=27.78243]
22: 198957 Babel_(newspaper)
    [pem_pos=10, final_scores=0.22513758,
final_local_scores=0.37678003, pem=-3.015456,
similarity=31.7321]
23: 7515928 Iraq
    [pem_pos=0, final_scores=0.68236506,
final_local_scores=0.62884873, pem=-0.25379938,
similarity=24.236013]
24: 7515964 Jordan
    [pem_pos=0, final_scores=0.58160573,
final_local_scores=0.5364816, pem=-0.53647494,
similarity=23.016829]
```

Sample: b'296_Clinton'

[1 1 Clinton 1] campaign busy making " news " . Laurence McQuillan [2 2 WASHINGTON 2] 1996-08-24 President [3 3 Bill Clinton 3] has served notice he intends to be busy " making news " -- or at least doing things that look and sound like it in a campaign year . With [4 4 Democrats 4] gathering in [5 5 Chicago 5] to start a convention on Monday to nominate him for a second term as president , [6 6 Clinton 6] plans a steady parade of events designed to highlight his leadership and dim the glow of the just-concluded [7 7 Republican 7] conclave that gave a boost to rival [8 8 Bob Dole 8] . After a week of carefully orchestrated events signing into law bills passed by the Republican-controlled [9 9 Congress 9] , [10 10 Clinton 10] used his Saturday radio address to the nation to proudly " announce " a development in the war on crime . " Sixty days ago I directed the attorney general to draw up a plan for a national registry of sex offenders , " [1 1 11 Clinton 11] said . " That plan has now reached my desk . " " Today I am pleased to announce that we are following through on our commitment to keep track of these criminals , not just in a single state but wherever they go , " he said . Actually , creation of such a registry was underway without [1 2 12 Clinton 12] lifting a finger . Attorney General [13 13 Janet Reno 13] 's report -- all nine pages of it , including footnotes -- offers only the interim services of the [14 14 FB I 14] until a formal registry on sex offenders has been established . Two months ago , [15 15 Clinton 15] announced he wanted an interim effort established . Now , 60 days later , he had a chance to talk about it again . It is an example of [16 16 Clinton 16] 's strategic planning as he heads into the stretch drive for the Nov. 5 presidential election . Such things do not happen by chance in the [17 17 Clinton 17] [18 18 White House 18] , they are part of his political chess game . In the past week [19 19 Clinton 19] signed into law an increase in the minimum wage , a bill that makes it easier for someone with a pre-existing health problem to change jobs , and sweeping changes overhauling the nation 's welfare system . " [20 20 America 20] can look back

on a week of remarkable achievement , " [21 21 Clinton 21] said , without even a passing reference to the [22 22 Republican 22] majority in the House and Senate . " [23 23 America 23] is on the right track offering more opportunity , demanding more responsibility , building a stronger community , the sense of shared values and stronger families , " he said in striking the theme of his coming week . According to a senior campaign official , [24 24 Clinton 24] " will be making a lot of news in the coming week -- something different each day . " [25 25 Clinton 25] departs on Sunday on a four-day train trip through [26 26 West Virginia 26] , [27 27 Kentucky 27] , [28 28 Ohio 28] , [29 29 Michigan 29] and [30 30 Indiana 30] while fellow [31 31 Democrats 31] are gathered in [32 32 Chicago 32] . " He 'll make news during the day ... and then at night the attention will go to the convention , " said the campaign official . " We think it 'll work really well . " Although officials decline to say just what each day 's " news " will be , the intent is to put a focus on [33 33 Clinton 33] himself and not just those attending the party 's convention . Each day of the trip will have a late start , so that network television correspondents will be able to do live reports for morning programmes . Campaign officials then hope the day 's " news " event will be showcased on evening television news shows as a lead-in for that night 's convention programme . " We 'll be concentrating ... on supporting the president as he is on the trip and making significant public policy statements related to some of his plans for the future , " said [34 34 White House 34] Press Secretary [35 35 Mike McCurry 35] . McCurry said that when [36 36 Clinton 36] delivers his acceptance address on Thursday night to fellow [37 37 Democrats 37] and a national television audience , he will " set out a road map " for the nation 's future -- one the president hopes guides him back to the [38 38 White House 38] for four more years .

- 1: 3356 Bill_Clinton
- 2: 108956 Washington,_D.C.
- 3: 3356 Bill_Clinton
- 4: 5043544 Democratic_Party_(United_States)
- 5: 6886 Chicago
- 6: 3356 Bill_Clinton
- 7: 32070 Republican_Party_(United_States)
- 8: 55846 Bob_Dole
- 9: 31756 United_States_Congress
- 10: 3356 Bill_Clinton
- 11: 3356 Bill_Clinton
- 12: 3356 Bill_Clinton
- 13: 16099 Janet_Reno
- 14: 11127 Federal_Bureau_of_Investigation
- 15: 3356 Bill_Clinton
- 16: 3356 Bill_Clinton
- 17: 3356 Bill_Clinton
- 18: 33057 White_House

```
19: 3356 Bill_Clinton
20: 3434750 United_States
21: 3356 Bill_Clinton
22: 32070 Republican_Party_(United_States)
23: 3434750 United_States
24: 3356 Bill_Clinton
25: 3356 Bill_Clinton
26: 32905 West_Virginia
27: 16846 Kentucky
28: 22199 Ohio
29: 18859 Michigan
30: 21883857 Indiana
31: 5043544 Democratic_Party_(United_States)
32: 6886 Chicago
33: 3356 Bill_Clinton
34: 33057 White_House
35: 284483 Mike_McCurry_(press_secretary)
36: 3356 Bill_Clinton
37: 5043544 Democratic_Party_(United_States)
38: 33057 White_House
1: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.809753,
final_local_scores=0.76777947, pem=0.0, similarity=27.001783]
2: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.5651795,
final_local_scores=0.55101746, pem=-1.7496481,
similarity=30.048033]
3: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.61977303,
final_local_scores=0.5597679, pem=0.0, similarity=20.794304]
4: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.47342575,
final_local_scores=0.4626149, pem=-0.18007085,
similarity=18.874317]
5: 6886 Chicago
    [pem_pos=0, final_scores=0.6889682,
final_local_scores=0.72304267, pem=-0.20658566,
similarity=26.790192]
6: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.9030705,
final_local_scores=0.8697147, pem=0.0, similarity=30.043734]
7: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.46105832,
final_local_scores=0.4478811, pem=-0.1310191,
similarity=18.167881]
8: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.69276214,
final_local_scores=0.704315, pem=0.0, similarity=25.107878]
9: 31756 United_States_Congress
    [pem_pos=0, final_scores=0.42835903,
final_local_scores=0.40359968, pem=-0.3349529,
similarity=17.95546]
```

```
10: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.85311204,
final_local_scores=0.8151425, pem=0.0, similarity=28.415192]
11: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.8619799,
final_local_scores=0.8248294, pem=0.0, similarity=28.704268]
12: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.887147,
final_local_scores=0.85232055, pem=0.0, similarity=29.524658]
13: 16099 Janet_Reno
    [pem_pos=0, final_scores=0.46492475,
final_local_scores=0.47605646, pem=0.0, similarity=18.296186]
14: 11127 Federal_Bureau_of_Investigation
    [pem_pos=0, final_scores=0.4427343,
final_local_scores=0.4705745, pem=0.0, similarity=18.132595]
15: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.8190399,
final_local_scores=0.7779239, pem=0.0, similarity=27.304514]
16: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.8700705,
final_local_scores=0.8336672, pem=0.0, similarity=28.968006]
17: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.83578175,
final_local_scores=0.7962119, pem=0.0, similarity=27.850264]
18: 33057 White_House
    [pem_pos=0, final_scores=0.44147277,
final_local_scores=0.40201306, pem=-0.0056634876,
similarity=16.117384]
19: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.86215574,
final_local_scores=0.8250215, pem=0.0, similarity=28.710001]
20: 3434750 United_States
    [pem_pos=0, final_scores=0.4267369,
final_local_scores=0.43512774, pem=-0.78483206,
similarity=21.34283]
21: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.87279445,
final_local_scores=0.8366427, pem=0.0, similarity=29.056799]
22: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.4213729,
final_local_scores=0.40453082, pem=-0.1310191,
similarity=16.874222]
23: 3434750 United_States
    [pem_pos=0, final_scores=0.50309473,
final_local_scores=0.51853716, pem=-0.78483206,
similarity=23.831934]
24: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.81374484,
final_local_scores=0.7721399, pem=0.0, similarity=27.131908]
25: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.8221255,
final_local_scores=0.7812945, pem=0.0, similarity=27.4051]
```

```
26: 32905 West_Virginia
    [pem_pos=0, final_scores=0.49096465,
final_local_scores=0.5327454, pem=-0.30025622,
similarity=21.620739]
27: 16846 Kentucky
    [pem_pos=0, final_scores=0.43043447,
final_local_scores=0.47587824, pem=-0.28237545,
similarity=19.826471]
28: 22199 Ohio
    [pem_pos=0, final_scores=0.63726074,
final_local_scores=0.6768818, pem=-0.19147073,
similarity=25.330463]
29: 18859 Michigan
    [pem_pos=0, final_scores=0.6660045,
final_local_scores=0.70907575, pem=-0.24788691,
similarity=26.597994]
30: 21883857 Indiana
    [pem_pos=0, final_scores=0.5900663,
final_local_scores=0.6163533, pem=-0.27230176,
similarity=23.963743]
31: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.48083287,
final_local_scores=0.47070605, pem=-0.18007085,
similarity=19.115774]
32: 6886 Chicago
    [pem_pos=0, final_scores=0.575867,
final_local_scores=0.5994968, pem=-0.20658566,
similarity=23.103338]
33: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.8790297,
final_local_scores=0.8434537, pem=0.0, similarity=29.260054]
34: 33057 White_House
    [pem_pos=1, final_scores=0.41234803,
final_local_scores=0.37019873, pem=-0.0056634876,
similarity=15.16798]
35: 284483 Mike_McCurry_(press_secretary)
    [pem_pos=0, final_scores=0.4490044,
final_local_scores=0.50378734, pem=0.0, similarity=19.123732]
36: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.87167245,
final_local_scores=0.835417, pem=0.0, similarity=29.020222]
37: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.4698698,
final_local_scores=0.45873052, pem=-0.18007085,
similarity=18.7584]
38: 33057 White_House
    [pem_pos=1, final_scores=0.475439,
final_local_scores=0.439116, pem=-0.0056634876,
similarity=17.22461]
```

Sample: b'297_Hurricane'

Hurricane expected to veer north of [1 1 Caribbean 1] . [2 2

MIAMI 2] 1996-08-24 Hurricane [3 3 Edouard 3] grew stronger on Saturday as it swirled across the Atlantic Ocean , but forecasters at the [4 4 National Hurricane Center 4] said the storm would likely swing north and miss the [5 5 Caribbean 5] . " [6 6 Edouard 6] is getting stronger and stronger , and it already has winds of 105 mph (185 kph) , " said hurricane forecaster [7 7 Lixion Avila 7] . " But the good news is that all our computer models indicate [8 8 Edouard 8] is going to turn to the west-northwest on Sunday and miss the islands , " he added . At 11 a.m. EDT (1500 [9 9 GMT 9]) , [10 10 Edouard 10] was 1,130 miles east of the Lesser Antilles and moving west at 14 mph (25 kph) . Its exact position was latitude 14.5 north , longitude 44.2 west .

```
1: 18956035 Caribbean
2: 53846 Miami
3: 3589067 Hurricane_Edouard_(1996)
4: 255317 National_Hurricane_Center
5: 18956035 Caribbean
6: 3589067 Hurricane_Edouard_(1996)
7: 3343855 Lixion_Avila
8: 3589067 Hurricane_Edouard_(1996)
9: 12701 Greenwich_Mean_Time
10: 3589067 Hurricane_Edouard_(1996)
1: 18956035 Caribbean
    [pem_pos=0, final_scores=0.4062382,
final_local_scores=0.44202477, pem=-0.09554641,
similarity=17.80021]
2: 53846 Miami
    [pem_pos=0, final_scores=0.44394267,
final_local_scores=0.51223075, pem=-0.35396606,
similarity=21.300623]
3: 3589067 Hurricane_Edouard_(1996)
    [pem_pos=0, final_scores=0.610238,
final_local_scores=0.5998388, pem=-1.0413741,
similarity=27.653254]
4: 255317 National_Hurricane_Center
    [pem_pos=0, final_scores=0.40539426,
final_local_scores=0.41085684, pem=0.0, similarity=16.3505]
5: 18956035 Caribbean
    [pem_pos=0, final_scores=0.32550615,
final_local_scores=0.3538373, pem=-0.09554641,
similarity=15.168521]
6: 3589067 Hurricane_Edouard_(1996)
    [pem_pos=0, final_scores=0.5164681,
final_local_scores=0.4974094, pem=-1.0413741,
similarity=24.596556]
7: 3343855 Lixion_Avila
    [pem_pos=0, final_scores=0.30910546,
final_local_scores=0.45674253, pem=0.0, similarity=17.71982]
8: 3589067 Hurricane_Edouard_(1996)
    [pem_pos=0, final_scores=0.5547616,
final_local_scores=0.53923917, pem=-1.0413741,
```

```
similarity=25.84484]
9: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.66035914,
final_local_scores=0.7876067, pem=0.0, similarity=27.593468]
10: 3589067 Hurricane_Eduard_(1996)
    [pem_pos=0, final_scores=0.3908115,
final_local_scores=0.36014867, pem=-1.0413741,
similarity=20.500422]
```

Sample: b'298_France'

```
[1 1 France 1] hands suspected [2 2 ETA 2] member to [3 3 Sp
ain 3] . [4 4 PARIS 4] 1996-08-24 [5 5 France 5] on Saturday
handed a suspected member of the [1 1 Basque 1] separatist
group [6 6 ETA 6] to [7 7 Spanish 7] authorities , [8 8 French
8] [9 9 Interior Ministry 9] officials said . Ignacio
Olascoaga Mugica , who had just ended a prison sentence in [10
10 France 10] , is suspected of having taken part in several
guerrilla attacks in [11 11 Spain 11] . [12 12 ETA 12] ( [13 1
3 Basque Homeland and Freedom 13] ) has killed about 800
people in its campaign for an independent [14 14 Basque 14]
state since the 1960s .
```

```
1: 5843419 France
2: 9926 ETA
3: 26667 Spain
4: 22989 Paris
5: 5843419 France
6: 9926 ETA
7: 26667 Spain
8: 5843419 France
9: 354620 Interior_ministry
10: 5843419 France
11: 26667 Spain
12: 9926 ETA
13: 9926 ETA
14: 683036 Basque_Country_(greater_region)
1: 683036 Basque_Country_(greater_region)
    [winner: 20514897 Basque_Country_(autonomous_community),
pem_pos=2, final_scores=0.1788358,
final_local_scores=0.14702052, pem=-2.0219665,
similarity=19.472866|
     gt: 683036 Basque_Country_(greater_region), pem_pos=3,
final_scores=0.11956404, final_local_scores=0.07250351,
pem=-2.163085, similarity=18.016554 ]
1: 5843419 France
    [pem_pos=0, final_scores=0.71216446,
final_local_scores=0.70425093, pem=-0.16240834,
similarity=25.989166]
2: 9926 ETA
    [pem_pos=0, final_scores=0.47259146,
final_local_scores=0.48304653, pem=-0.10190937,
similarity=19.058985]
3: 26667 Spain
```

```
[pem_pos=0, final_scores=0.59193367,
final_local_scores=0.5633904, pem=-0.18857051,
similarity=21.927881]
4: 22989 Paris
    [pem_pos=0, final_scores=0.49903864,
final_local_scores=0.5147282, pem=-0.115301244,
similarity=20.077255]
5: 5843419 France
    [pem_pos=0, final_scores=0.6584906,
final_local_scores=0.6456204, pem=-0.16240834,
similarity=24.239515]
6: 9926 ETA
    [pem_pos=0, final_scores=0.50038344,
final_local_scores=0.513405, pem=-0.10190937,
similarity=19.964943]
7: 26667 Spain
    [pem_pos=1, final_scores=0.70837754,
final_local_scores=0.6905876, pem=-1.2425895,
similarity=31.455618]
8: 5843419 France
    [pem_pos=0, final_scores=0.642429,
final_local_scores=0.62807554, pem=-0.80223274,
similarity=27.195406]
9: 354620 Interior_ministry
    [pem_pos=0, final_scores=0.11724745,
final_local_scores=0.16268638, pem=-0.57122487,
similarity=12.051009]
10: 5843419 France
    [pem_pos=0, final_scores=0.7289942,
final_local_scores=0.72263485, pem=-0.16240834,
similarity=26.537779]
11: 26667 Spain
    [pem_pos=0, final_scores=0.6359417,
final_local_scores=0.6114624, pem=-0.18857051,
similarity=23.362446]
12: 9926 ETA
    [pem_pos=0, final_scores=0.50043154,
final_local_scores=0.5134576, pem=-0.10190937,
similarity=19.96651]
13: 9926 ETA
    [pem_pos=0, final_scores=0.64787775,
final_local_scores=0.6745202, pem=0.0, similarity=24.21874]
14: 683036 Basque_Country_(greater_region)
    [pem_pos=3, final_scores=0.12164636,
final_local_scores=0.07477814, pem=-2.163085,
similarity=18.084435]
1: 20514897 Basque_Country_(autonomous_community)
    [pem_pos=2, final_scores=0.1788358,
final_local_scores=0.14702052, pem=-2.0219665,
similarity=19.472866]
```

Sample: b'299_German'

[1 1 German 1] troops to remain in Bosnia for 1997--Ruehe . [
2 2 BONN 2] 1996-08-24 Defence Minister [3 3 Volker Ruehe 3]
said that [4 4 German 4] troops would stay on in Bosnia next
year as part of an international force to ensure the
establishment of peace , a newspaper reported on Saturday .
The current NATO-led peace force ([5 5 IFOR 5]) in Bosnia
is due to return home at the end of the year . But Ruehe told
[6 6 Bild am Sonntag 6] in an interview that a " new and
different mandate " for the troops would be agreed on for
next year after the current mandate expires in December . "
After the (Bosnian) elections (on September 14) the
troops will start being reduced from the beginning of October
from 60,000 to about 20,000 . A completely new and different
mandate will be agreed for next year , " Ruehe said . " The
defence ministers will begin negotiations for this at the
beginning of September at a [7 7 NATO 7] meeting , " he told
the newspaper in an interview , excerpts of which were
released ahead of publication on Sunday . " But we must avoid
giving the impression this peace deployment in former [8 8 Yu
goslavia 8] is being perceived in the long run as an
occupation . On the other hand we must prevent any return of
war and massacres " he said .

1: 11867 Germany
2: 3295 Bonn
3: 747568 Volker_Rühe
4: 11867 Germany
5: 374695 Implementation_Force
6: 237183 Bild
7: 21133 NATO
8: 34244 Yugoslavia
1: 11867 Germany
 [pem_pos=1, final_scores=0.76920944,
final_local_scores=0.7690745, pem=-1.0465823,
similarity=32.731907]
2: 3295 Bonn
 [pem_pos=0, final_scores=0.4632135,
final_local_scores=0.4490509, pem=0.0, similarity=17.490288]
3: 747568 Volker_Rühe
 [pem_pos=0, final_scores=0.41628397,
final_local_scores=0.52123284, pem=0.0, similarity=19.64434]
4: 11867 Germany
 [pem_pos=1, final_scores=0.7801822,
final_local_scores=0.7810605, pem=-1.0465823,
similarity=33.089596]
5: 374695 Implementation_Force
 [pem_pos=0, final_scores=0.47327858,
final_local_scores=0.55764055, pem=0.0, similarity=20.73082]
6: 237183 Bild
 [pem_pos=0, final_scores=0.42469913,
final_local_scores=0.5085533, pem=0.0, similarity=19.265959]
7: 21133 NATO
 [pem_pos=0, final_scores=0.3968528,

```
final_local_scores=0.40717703, pem=0.0, similarity=16.240688]
8: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.44637978,
final_local_scores=0.4317708, pem=-0.6536551,
similarity=20.529293]
```

Sample: b'300_Italian'

[1 1 Italian 1] comics hope independence-joke 's on [2 2 Bos
si 2] . [3 3 ORVIETO 3] , [4 4 Italy 4] 1996-08-24 A group of
[5 5 Italian 5] comics hope the joke will be on separatist
leader [6 6 Umberto Bossi 6] next month when they lead the
ancient [7 7 Etruscan 7] town of [8 8 Orvieto 8] in a mock
split from [9 9 Rome 9] . [10 10 Orvieto 10] mayor Stefano
Cimicchi said the comics , including popular actor [11 11 Rob
erto Benigni 11] , would declare [12 12 Orvieto 12] " capital
of [13 13 Etruria 13] " on September 15 -- the day [14 14 Bos
si 14] plans a march across the north in favour of
independence from [15 15 Rome 15] . " We will then proceed
with the annexation of [16 16 Sardinia 16] , [17 17 Corsica 17
] and [18 18 Cyprus 18] , " Cimicchi told reporters on
Saturday . He said the city council would be " ironically
present " when the comics made their proclamation on the same
day [19 19 Bossi 19] has threatened to declare the birth of
Padania , the name he has given to northern [20 20 Italy 20] .
[21 21 Orvieto 21] , located in [22 22 Umbria 22] between [23
23 Rome 23] and [24 24 Florence 24] , was once the capital of
[25 25 Etruria 25] , an ancient federation of 12 [26 26 Etrus
can 26] towns . " We want to pop some air out of this balloon
of tension that has been blown up around September 15 , "
Cimicchi said . " We want to help turn down the rhetoric in a
country that borders former [27 27 Yugoslavia 27] yet in
which people are still talking about secession , " he added .
[28 28 Bossi 28] has intensified his separatist rhetoric
since his [29 29 Northern League 29] party 's good showing in
last April 's general election , when it took 10 percent of
the vote nationally . He has recently dropped a drive for
federalism , saying secession from [30 30 Rome 30] 's
wasteful and centralised bureaucracy is the only solution for
northerners .

```
1: 14532 Italy
2: 842061 Umberto_Bossi
3: 221674 Orvieto
4: 14532 Italy
5: 14532 Italy
6: 842061 Umberto_Bossi
7: 72845 Etruria
8: 221674 Orvieto
9: 25458 Rome
10: 221674 Orvieto
11: 46344 Roberto_Benigni
12: 221674 Orvieto
13: 72845 Etruria
```

```
14: 842061 Umberto_Bossi
15: 25458 Rome
16: 29376 Sardinia
17: 5714828 Corsica
18: 5593 Cyprus
19: 842061 Umberto_Bossi
20: 14532 Italy
21: 221674 Orvieto
22: 51590 Umbria
23: 25458 Rome
24: 11525 Florence
25: 72845 Etruria
26: 72845 Etruria
27: 34244 Yugoslavia
28: 842061 Umberto_Bossi
29: 357391 Lega_Nord
30: 25458 Rome
1: 14532 Italy
    [pem_pos=0, final_scores=0.7838996,
final_local_scores=0.769777, pem=-0.75948656,
similarity=31.191599]
2: 842061 Umberto_Bossi
    [pem_pos=0, final_scores=0.5161066,
final_local_scores=0.5481308, pem=0.0, similarity=20.44703]
3: 221674 Orvieto
    [pem_pos=0, final_scores=0.38293177,
final_local_scores=0.35213724, pem=0.0, similarity=14.59819]
4: 14532 Italy
    [pem_pos=0, final_scores=0.76716703,
final_local_scores=0.7514992, pem=-0.20612453,
similarity=27.636885]
5: 14532 Italy
    [pem_pos=0, final_scores=0.7320664,
final_local_scores=0.71315706, pem=-0.75948656,
similarity=29.501947]
6: 842061 Umberto_Bossi
    [pem_pos=0, final_scores=0.60190105,
final_local_scores=0.64184827, pem=0.0, similarity=23.243746]
7: 72845 Etruria
    [pem_pos=3, final_scores=0.27988046,
final_local_scores=0.27940828, pem=-1.8814839,
similarity=22.659615]
8: 221674 Orvieto
    [pem_pos=0, final_scores=0.44297355,
final_local_scores=0.41772377, pem=0.0, similarity=16.555424]
9: 25458 Rome
    [pem_pos=0, final_scores=0.5518535,
final_local_scores=0.51019704, pem=-0.2517334,
similarity=20.683975]
10: 221674 Orvieto
    [pem_pos=0, final_scores=0.42964166,
final_local_scores=0.40316075, pem=0.0, similarity=16.120834]
```

```
11: 46344 Roberto_Benigni
    [pem_pos=0, final_scores=0.36164594,
final_local_scores=0.443743, pem=0.0, similarity=17.331888]
12: 221674 Orvieto
    [pem_pos=0, final_scores=0.4889924,
final_local_scores=0.46799242, pem=0.0, similarity=18.05554]
13: 72845 Etruria
    [pem_pos=0, final_scores=0.47183305,
final_local_scores=0.48908728, pem=-0.07673108,
similarity=19.10233]
14: 842061 Umberto_Bossi
    [pem_pos=0, final_scores=0.597301,
final_local_scores=0.63682336, pem=0.0, similarity=23.093792]
15: 25458 Rome
    [pem_pos=0, final_scores=0.47902524,
final_local_scores=0.43064326, pem=-0.2517334,
similarity=18.30993]
16: 29376 Sardinia
    [pem_pos=0, final_scores=0.55948514,
final_local_scores=0.55127776, pem=0.0, similarity=20.540941]
17: 5714828 Corsica
    [pem_pos=0, final_scores=0.45910525,
final_local_scores=0.47134542, pem=-0.040366642,
similarity=18.375122]
18: 5593 Cyprus
    [pem_pos=0, final_scores=0.56215984,
final_local_scores=0.6032787, pem=-0.3752463,
similarity=24.1334]
19: 842061 Umberto_Bossi
    [pem_pos=0, final_scores=0.56029224,
final_local_scores=0.5963969, pem=0.0, similarity=21.887386]
20: 14532 Italy
    [pem_pos=0, final_scores=0.7124507,
final_local_scores=0.6917299, pem=-0.20612453,
similarity=25.85325]
21: 221674 Orvieto
    [pem_pos=0, final_scores=0.48294502,
final_local_scores=0.46138656, pem=0.0, similarity=17.858408]
22: 51590 Umbria
    [pem_pos=0, final_scores=0.45503682,
final_local_scores=0.42133147, pem=0.0, similarity=16.663084]
23: 25458 Rome
    [pem_pos=0, final_scores=0.4674995,
final_local_scores=0.41805315, pem=-0.2517334,
similarity=17.934217]
24: 11525 Florence
    [pem_pos=0, final_scores=0.49094844,
final_local_scores=0.45723307, pem=-0.32229787,
similarity=19.487165]
25: 72845 Etruria
    [pem_pos=0, final_scores=0.45201343,
final_local_scores=0.4674374, pem=-0.07673108,
```

```
similarity=18.456253]
26: 72845 Etruria
    [pem_pos=3, final_scores=0.2756588,
final_local_scores=0.27479672, pem=-1.8814839,
similarity=22.521997]
27: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.3926614,
final_local_scores=0.4119135, pem=-0.6536551,
similarity=19.936712]
28: 842061 Umberto_Bossi
    [pem_pos=0, final_scores=0.65967375,
final_local_scores=0.7049561, pem=0.0, similarity=25.12701]
29: 357391 Lega_Nord
    [pem_pos=1, final_scores=0.40182585,
final_local_scores=0.45026678, pem=-1.1138831,
similarity=23.584042]
30: 25458 Rome
    [pem_pos=0, final_scores=0.56575406,
final_local_scores=0.5253813, pem=-0.2517334,
similarity=21.137104]
```

Sample: b'301_Italian'

[1 1 Italian 1] farmer says he mutilated four women . [2 2 VE RONA 2] , [3 3 Italy 3] 1996-08-24 An [4 4 Italian 4] farmer accused of multiple homicide has confessed to mutilating the bodies of four women after having sex with them , the [5 5 It alian 5] news agency [6 6 ANSA 6] reported on Saturday . It quoted the lawyer of Gianfranco Stevanin as saying the 35-year-old farmer confessed on Friday to a [7 7 Verona 7] magistrate that he had killed and mutilated the women . [8 8 A NSA 8] said Stevanin was unable to recall how he had killed the women , remembering only that he had found them " lifeless in his arms " after having sadomasochistic sex with them . Stevanin , arrested in 1994 and jailed for three years for assaulting an [9 9 Austrian 9] prostitute , is accused of murdering five women , three of whose bodies were found near his villa outside [10 10 Verona 10] between July and December 1995 . Two of the corpses were identified but not the third , found headless and decomposed in a sack in a nearby canal . Lawyer Cesare dal Maso told [11 11 ANSA 11] that Stevanin confessed to beheading and dumping the body of a fourth woman in the nearby [12 12 Adige 12] river . Dal Maso declined to comment on the alleged fifth murder , saying only that " the interrogations are not over yet " with investigators . It said investigators believed Stevanin had suffocated them by putting plastic bags on their heads . Stevanin was first sentenced for assault but investigators began digging in the garden of his villa after the first body was found by a passer-by .

```
1: 14532 Italy
2: 263991 Verona
3: 14532 Italy
```

```
4: 14532 Italy
5: 14532 Italy
6: 10088602 Agenzia_Nazionale_Stampa_Associata
7: 263991 Verona
8: 10088602 Agenzia_Nazionale_Stampa_Associata
9: 26964606 Austria
10: 263991 Verona
11: 10088602 Agenzia_Nazionale_Stampa_Associata
12: 592728 Adige
1: 14532 Italy
    [pem_pos=1, final_scores=0.71497554,
final_local_scores=0.68485266, pem=-0.75948656,
similarity=28.657288]
2: 263991 Verona
    [pem_pos=0, final_scores=0.51971275,
final_local_scores=0.48566407, pem=-0.4457646,
similarity=21.007034]
3: 14532 Italy
    [pem_pos=0, final_scores=0.7495094,
final_local_scores=0.72257566, pem=-0.20612453,
similarity=26.77375]
4: 14532 Italy
    [pem_pos=0, final_scores=0.7575864,
final_local_scores=0.73139864, pem=-0.75948656,
similarity=30.046312]
5: 14532 Italy
    [pem_pos=1, final_scores=0.6665564,
final_local_scores=0.63196224, pem=-0.75948656,
similarity=27.078932]
6: 10088602 Agenzia_Nazionale_Stampa_Associata
    [pem_pos=0, final_scores=0.4633662,
final_local_scores=0.5060314, pem=0.0, similarity=19.190699]
7: 263991 Verona
    [pem_pos=0, final_scores=0.60197586,
final_local_scores=0.57552403, pem=-0.4457646,
similarity=23.688635]
8: 10088602 Agenzia_Nazionale_Stampa_Associata
    [pem_pos=0, final_scores=0.4052347,
final_local_scores=0.44253153, pem=0.0, similarity=17.295736]
9: 26964606 Austria
    [pem_pos=0, final_scores=0.7229753,
final_local_scores=0.7506332, pem=-0.40674555,
similarity=28.70205]
10: 263991 Verona
    [pem_pos=0, final_scores=0.45455247,
final_local_scores=0.4144864, pem=-0.4457646,
similarity=18.882952]
11: 10088602 Agenzia_Nazionale_Stampa_Associata
    [pem_pos=0, final_scores=0.4572758,
final_local_scores=0.49937856, pem=0.0, similarity=18.992165]
12: 592728 Adige
    [pem_pos=0, final_scores=0.35956544,
```

final_local_scores=0.40885848, pem=0.0, similarity=16.290865]

Sample: b'302_Belgium'

[1 1 Belgium 1] asks how paedophile suspect eluded police . Jeremy Lovell [2 2 BRUSSELS 2] 1996-08-24 [3 3 Belgian 3] police searched two more houses on Saturday for bodies in a child-sex scandal of murder , kidnapping and pornography that has sent a shockwave of revulsion throughout [4 4 Europe 4] . Recriminations built up over how the scandal 's central figure , convicted child rapist [5 5 Marc Dutroux 5] , managed to prey on children unhindered for so long . In just over a week two young girls have been found dead , from starvation , two have been freed from a dungeon-like secret compartment and an international hunt has started for at least two others . On Saturday investigators with dogs trained to find bodies searched one house at [6 6 Ransart 6] and one at [7 7 Mont-sur-Marchienne 7] -- both suburbs of the southern city of [8 8 Charleroi 8] . Both houses are owned by [9 9 Dutroux 9] . [10 10 Belgian 10] media speculated that [11 11 Dutroux 11] , charged with abduction and illegal imprisonment of children , must have had high level protection to molest youngsters . They put forward no proof to support the speculation , but seized on a comment by chief prosecutor Michel Bourlet on [12 12 Belgian 12] television on Friday night that he would chase down everyone involved in the case " if I am allowed to " . Bourlet said between 300 and 400 paedophile porn video tapes had been seized , some of which featured [13 13 Dutroux 13] . [14 14 Dutroux 14] was charged a week ago after police rescued two young girls from a concrete dungeon in the basement of one of the six houses he owns in and around [15 15 Charleroi 15] . Just a day later the national euphoria at the rescue turned to disgust as [16 16 Dutroux 16] led police to the bodies of two eight-year-old girls in another of his houses . Julie Lejeune and Melissa Russo , had been kidnapped in June last year . [17 17 Dutroux 17] said they starved to death nine months later . He also admitted kidnapping two other girls , An Marchal and Eefje Lambrecks , a year ago . The fate of the girls is unknown , but there has been speculation they were sold into prostitution in [18 18 Slovakia 18] or the [19 19 Czech Republic 19] where [20 20 Dutroux 20] was a frequent visitor . [21 21 Belgian 21] police have visited [22 22 Bratislava 22] and will visit [23 23 Prague 23] . Five other people have been arrested including [24 24 Dutroux 24] 's second wife Michelle Martin , charged as an accomplice . The others have been charged with abduction and illegal imprisonment of children or are suspected of criminal association . [25 25 Dutch 25] police are also holding a 74-year old Dutchman in connection with the disappearance of An and Eefje , although a spokesman said no direct link had yet been established . At least part of the speculation in the [26 26 Belgian 26] media of high-level protection for [27 27 Dutroux 27] and his accomplices

is based on leaked documents cataloguing a high degree of police bungling , incompetence and indifference . Among the revelations are the fact that the gendarmerie was running a surveillance operation codenamed " Othello " against [28 28 Du troux 28] in 1995 -- when both Julie and Melissa and An and Eefje were kidnapped . They show that the gendarmes were aware that [29 29 Dutroux 29] was building cells in some of his houses for holding children , yet this information was either not passed on to other police forces searching for the missing girls or was overlooked when it was . They also show that police investigating a theft visited [30 30 Dutroux 30] late last year at the house where Julie and Melissa were being held but accepted his word that the children 's cries they could hear came from neighbours . Justice Minister [31 31 Stefaan De Clerck 31] has admitted that mistakes were made and ordered an inquiry at the same time as stressing there were no indications of a cover-up . There is also widespread disbelief that no one appeared to question how [32 32 Dutroux 32] , an unemployed father of three with no visible means of support , managed to own six houses .

1: 3343 Belgium
2: 3708 Brussels
3: 3343 Belgium
4: 9239 Europe
5: 154694 Marc_Dutroux
6: 17703371 Ransart,_Belgium
7: 17703335 Mont-sur-Marchienne
8: 175228 Charleroi
9: 154694 Marc_Dutroux
10: 3343 Belgium
11: 154694 Marc_Dutroux
12: 3343 Belgium
13: 154694 Marc_Dutroux
14: 154694 Marc_Dutroux
15: 175228 Charleroi
16: 154694 Marc_Dutroux
17: 154694 Marc_Dutroux
18: 26830 Slovakia
19: 5321 Czech_Republic
20: 154694 Marc_Dutroux
21: 3343 Belgium
22: 18933194 Bratislava
23: 23844 Prague
24: 154694 Marc_Dutroux
25: 21148 Netherlands
26: 3343 Belgium
27: 154694 Marc_Dutroux
28: 154694 Marc_Dutroux
29: 154694 Marc_Dutroux
30: 154694 Marc_Dutroux
31: 1104698 Stefaan_De_Clerck
32: 154694 Marc_Dutroux

```
1: 3343 Belgium
   [pem_pos=0, final_scores=0.5478683,
final_local_scores=0.53955454, pem=-0.24180195,
similarity=21.506054]
2: 3708 Brussels
   [pem_pos=0, final_scores=0.5472718,
final_local_scores=0.5482727, pem=-0.025440974,
similarity=20.589617]
3: 3343 Belgium
   [pem_pos=0, final_scores=0.6636887,
final_local_scores=0.66607076, pem=-0.12403572,
similarity=24.641117]
4: 9239 Europe
   [pem_pos=0, final_scores=0.62125933,
final_local_scores=0.705699, pem=-0.086338006,
similarity=25.6187]
5: 154694 Marc_Dutroux
   [pem_pos=0, final_scores=0.8543733,
final_local_scores=0.86286026, pem=0.0, similarity=29.839184]
6: 17703371 Ransart,_Belgium
   [pem_pos=1, final_scores=0.49971497,
final_local_scores=0.65743244, pem=-0.6325226,
similarity=27.148563]
7: 17703335 Mont-sur-Marchienne
   [pem_pos=0, final_scores=0.21631126,
final_local_scores=0.3649565, pem=0.0, similarity=14.980742]
8: 175228 Charleroi
   [pem_pos=0, final_scores=0.38831526,
final_local_scores=0.39706784, pem=-0.2544922,
similarity=17.322977]
9: 154694 Marc_Dutroux
   [pem_pos=0, final_scores=0.6022964,
final_local_scores=0.5875046, pem=0.0, similarity=21.622023]
10: 3343 Belgium
    [pem_pos=0, final_scores=0.66512996,
final_local_scores=0.6676451, pem=-0.12403572,
similarity=24.688099]
11: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.7117209,
final_local_scores=0.70703423, pem=0.0, similarity=25.189024]
12: 3343 Belgium
    [pem_pos=0, final_scores=0.6316616,
final_local_scores=0.63108605, pem=-0.12403572,
similarity=23.597105]
13: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.56353706,
final_local_scores=0.54516584, pem=0.0, similarity=20.35855]
14: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.5985752,
final_local_scores=0.5834397, pem=0.0, similarity=21.500717]
15: 175228 Charleroi
    [pem_pos=0, final_scores=0.42460614,
```

```
final_local_scores=0.43671012, pem=-0.2544922,  
similarity=18.505981]  
16: 154694 Marc_Dutroux  
    [pem_pos=0, final_scores=0.6230115,  
final_local_scores=0.61013263, pem=0.0, similarity=22.297287]  
17: 154694 Marc_Dutroux  
    [pem_pos=0, final_scores=0.59040993,  
final_local_scores=0.57452035, pem=0.0, similarity=21.234547]  
18: 26830 Slovakia  
    [pem_pos=0, final_scores=0.52654034,  
final_local_scores=0.5761529, pem=-0.36602598,  
similarity=23.273771]  
19: 5321 Czech_Republic  
    [pem_pos=0, final_scores=0.37262923,  
final_local_scores=0.39818913, pem=-0.27794555,  
similarity=17.483982]  
20: 154694 Marc_Dutroux  
    [pem_pos=0, final_scores=0.7008396,  
final_local_scores=0.69514805, pem=0.0, similarity=24.834318]  
21: 3343 Belgium  
    [pem_pos=0, final_scores=0.6870416,  
final_local_scores=0.69158024, pem=-0.12403572,  
similarity=25.402372]  
22: 18933194 Bratislava  
    [pem_pos=0, final_scores=0.48120028,  
final_local_scores=0.52040595, pem=0.0, similarity=19.619665]  
23: 23844 Prague  
    [pem_pos=0, final_scores=0.3615679,  
final_local_scores=0.39094305, pem=-0.05392049,  
similarity=16.049461]  
24: 154694 Marc_Dutroux  
    [pem_pos=0, final_scores=0.7025717,  
final_local_scores=0.69704014, pem=0.0, similarity=24.890781]  
25: 21148 Netherlands  
    [pem_pos=0, final_scores=0.57327193,  
final_local_scores=0.59759337, pem=-0.7943452,  
similarity=26.242865]  
26: 3343 Belgium  
    [pem_pos=0, final_scores=0.6668872,  
final_local_scores=0.6695646, pem=-0.12403572,  
similarity=24.74538]  
27: 154694 Marc_Dutroux  
    [pem_pos=0, final_scores=0.74228317,  
final_local_scores=0.74041885, pem=0.0, similarity=26.185287]  
28: 154694 Marc_Dutroux  
    [pem_pos=0, final_scores=0.64809054,  
final_local_scores=0.63752776, pem=0.0, similarity=23.114813]  
29: 154694 Marc_Dutroux  
    [pem_pos=0, final_scores=0.6378853,  
final_local_scores=0.62638, pem=0.0, similarity=22.782143]  
30: 154694 Marc_Dutroux  
    [pem_pos=0, final_scores=0.6566474,
```

```
final_local_scores=0.6468747, pem=0.0, similarity=23.393745]
31: 1104698 Stefaan_De_Clerck
    [pem_pos=0, final_scores=0.4673857,
final_local_scores=0.60055286, pem=0.0, similarity=22.011408]
32: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.70637137,
final_local_scores=0.7011907, pem=0.0, similarity=25.014643]
```

Sample: b'303_Death'

Death toll of [1 1 Algeria 1] bomb put at seven-newspaper . [2 2 PARIS 2] 1996-08-24 An [3 3 Algerian 3] newspaper on Saturday put at seven -- two women and five children -- the death toll of a bomb blast in a market west of [4 4 Algiers 4] on Friday . [5 5 Algerian 5] security forces said on Friday three women and two children were killed and five people wounded when a home-made bomb exploded at a market in the coastal town of Bou Haroun , 65 km (40 miles) west of [6 6 Algiers 6] . The security forces also said a man carrying an explosive device also died after it went off prematurely . [7 7 El-Watan 7] paper said the blast killed seven -- a mother and her 25-year-old daughter , four young boys and a five-year-old girl . Several people were also wounded , it said . The explosion was the latest in series of bomb attacks in [8 8 Algeria 8] 's four-year-old civil strife . The government-appointed watchdog , Human Rights National Observatory , was quoted this month by local newspapers as saying about 1,400 civilians have died in bomb attacks blamed on [9 9 Moslem 9] rebels in the past two years . An estimated 50,000 [10 10 Algerians 10] and more than 110 foreigners have been killed in violence pitting [11 11 Moslem 11] rebels against government forces since early 1992 , when the authorities cancelled a general election in which radical [12 12 Islamists 12] had taken a commanding lead .

```
1: 358 Algeria
2: 22989 Paris
3: 358 Algeria
4: 1644 Algiers
5: 358 Algeria
6: 1644 Algiers
7: 4074715 El_Watan
8: 358 Algeria
9: 6037917 Islam
10: 358 Algeria
11: 6037917 Islam
12: 15012 Islamism
1: 358 Algeria
    [pem_pos=0, final_scores=0.72647125,
final_local_scores=0.6985976, pem=-0.33694586,
similarity=26.769623]
2: 22989 Paris
    [pem_pos=0, final_scores=0.47944456,
final_local_scores=0.51050824, pem=-0.115301244,
```

```
similarity=19.951323]
3: 358 Algeria
    [pem_pos=0, final_scores=0.5994355,
final_local_scores=0.5598304, pem=-0.20917186,
similarity=21.933678]
4: 1644 Algiers
    [pem_pos=0, final_scores=0.4101917,
final_local_scores=0.37482524, pem=-0.10460702,
similarity=15.844114]
5: 358 Algeria
    [pem_pos=0, final_scores=0.7028225,
final_local_scores=0.67276496, pem=-0.20917186,
similarity=25.30387]
6: 1644 Algiers
    [pem_pos=0, final_scores=0.39929956,
final_local_scores=0.36292726, pem=-0.10460702,
similarity=15.489056]
7: 4074715 El_Watan
    [pem_pos=0, final_scores=0.46414518,
final_local_scores=0.5817988, pem=0.0, similarity=21.45175]
8: 358 Algeria
    [pem_pos=0, final_scores=0.54838115,
final_local_scores=0.5040612, pem=-0.33694586,
similarity=20.96427]
9: 6037917 Islam
    [pem_pos=1, final_scores=0.46506697,
final_local_scores=0.5020512, pem=-1.2150909,
similarity=25.679773]
10: 358 Algeria
    [pem_pos=0, final_scores=0.2752558,
final_local_scores=0.20571327, pem=0.0, similarity=10.228607]
11: 6037917 Islam
    [pem_pos=1, final_scores=0.4253273,
final_local_scores=0.4586416, pem=-1.2150909,
similarity=24.384346]
12: 15012 Islamism
    [pem_pos=0, final_scores=0.13759722,
final_local_scores=0.14404827, pem=0.0, similarity=8.388401]
```

Sample: b'304_Malta'

```
[1 1 Malta 1] police seize cannabis among chilli sauce . [2 2
VALLETTA 2] 1996-08-24 Police in [3 3 Malta 3] said on
Saturday they had seized 7.5 tonnes of cannabis concealed in
a shipment of chilli sauce on its way from [4 4 Singapore 4]
to [5 5 Romania 5] . Police commissioner George Grech said
the cannabis was found on Friday packed in 500 boxes hidden
behind chilli sauce in a container that arrived at [6 6 Malta
Freeport 6] a week ago . The container was on its way to [7 7
Romania 7] via the former [8 8 Yugoslavia 8] from [9 9 Singap
ore 9] and was the biggest drugs haul in [10 10 Malta 10] ,
police said . No street value was given for the cannabis .
1: 19137 Malta
```

```
2: 57664 Valletta
3: 19137 Malta
4: 27318 Singapore
5: 25445 Romania
6: 28223454 Malta_Freeport
7: 25445 Romania
8: 34244 Yugoslavia
9: 27318 Singapore
10: 19137 Malta
1: 19137 Malta
    [pem_pos=0, final_scores=0.5704996,
final_local_scores=0.5264732, pem=-0.39222723,
similarity=21.933718]
2: 57664 Valletta
    [pem_pos=0, final_scores=0.438083,
final_local_scores=0.4431749, pem=-0.035661608,
similarity=17.50887]
3: 19137 Malta
    [pem_pos=0, final_scores=0.50437766,
final_local_scores=0.45424503, pem=-0.39222723,
similarity=19.778284]
4: 27318 Singapore
    [pem_pos=0, final_scores=0.46279866,
final_local_scores=0.45491463, pem=-0.1922004,
similarity=18.71049]
5: 25445 Romania
    [pem_pos=0, final_scores=0.6478949,
final_local_scores=0.660586, pem=-0.21241479,
similarity=24.958061]
6: 28223454 Malta_Freeport
    [pem_pos=0, final_scores=0.3279932,
final_local_scores=0.4691156, pem=0.0, similarity=18.089058]
7: 25445 Romania
    [pem_pos=0, final_scores=0.65706605,
final_local_scores=0.67060405, pem=-0.21241479,
similarity=25.257019]
8: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.4066221,
final_local_scores=0.3876168, pem=-0.6536551,
similarity=19.21165]
9: 27318 Singapore
    [pem_pos=0, final_scores=0.4307587,
final_local_scores=0.41991585, pem=-0.1922004,
similarity=17.666056]
10: 19137 Malta
    [pem_pos=0, final_scores=0.46058965,
final_local_scores=0.40641326, pem=-0.39222723,
similarity=18.350887]
```

Sample: b'305_Czech'
[1 1 Czech 1] coach in fatal crash in [2 2 Austria 2] . [3 3
VIENNA 3] 1996-08-24 A [4 4 Czech 4] coach crashed and burst

into flames on a southern [5 5 Austrian 5] motorway early on Saturday , killing one person and injuring 15 , police said . [6 6 Austrian 6] television said the coach , which was carrying 45 , was en route from the [7 7 Czech Republic 7] to [8 8 Italy 8] when the accident occurred near Steinberg , 200 km southwest of [9 9 Vienna 9] .

1: 5321 Czech_Republic
2: 26964606 Austria
3: 55866 Vienna
4: 5321 Czech_Republic
5: 26964606 Austria
6: 26964606 Austria
7: 5321 Czech_Republic
8: 14532 Italy
9: 55866 Vienna
1: 5321 Czech_Republic
 [pem_pos=1, final_scores=0.57766515,
final_local_scores=0.53452843, pem=-1.208016,
similarity=26.610487]
2: 26964606 Austria
 [pem_pos=0, final_scores=0.7327957,
final_local_scores=0.68483394, pem=-0.25262418,
similarity=25.900333]
3: 55866 Vienna
 [pem_pos=0, final_scores=0.6061829,
final_local_scores=0.583294, pem=-0.12517071,
similarity=22.177065]
4: 5321 Czech_Republic
 [pem_pos=1, final_scores=0.5841606,
final_local_scores=0.5416237, pem=-1.208016,
similarity=26.822224]
5: 26964606 Austria
 [pem_pos=0, final_scores=0.749346,
final_local_scores=0.7029126, pem=-0.40674555,
similarity=27.277973]
6: 26964606 Austria
 [pem_pos=0, final_scores=0.63598394,
final_local_scores=0.5790818, pem=-0.40674555,
similarity=23.582613]
7: 5321 Czech_Republic
 [pem_pos=0, final_scores=0.55819404,
final_local_scores=0.5132592, pem=-0.27794555,
similarity=20.917902]
8: 14532 Italy
 [pem_pos=0, final_scores=0.7713756,
final_local_scores=0.7833639, pem=-0.20612453,
similarity=28.587791]
9: 55866 Vienna
 [pem_pos=0, final_scores=0.5892711,
final_local_scores=0.56482047, pem=-0.12517071,
similarity=21.62578]

Sample: b'306_Most'

Most Spaniards back talks with [1 1 Basque 1] rebels--poll .
[2 2 MADRID 2] 1996-08-24 Most Spaniards would support
government talks with the illegal [3 3 Basque 3] separatist
group [4 4 ETA 4] if the rebels renounced violence
permanently , a survey published in daily [5 5 El Mundo 5] on
Saturday said . While 57 percent of the population supported
negotiations with [6 6 ETA 6] ([7 7 Basque Homeland and Free
dom 7]) , 30 percent opposed it , the survey by the state-
controlled Centre for Sociological Studies (CIS) found .
But 80 percent said [8 8 ETA 8] had shown little interest in
achieving peace in the [9 9 Basque 9] country when it offered
a one-week truce in July while continuing to hold prison
officer Jose Antonio Ortega Lara , kidnapped in January . The
problem of terrorism had neither worsened nor improved since
the conservative [10 10 Popular Party 10] (PP) came to
power in May , according to 56 percent of those questioned ,
while 22 percent said it had worsened . The survey questioned
2,496 people between July 17 and 21 and has a margin of error
of plus or minus two percent .

1: 20514897 Basque_Country_(autonomous_community)
2: 41188263 Madrid
3: 20514897 Basque_Country_(autonomous_community)
4: 9926 ETA
5: 539568 El_Mundo_(Spain)
6: 9926 ETA
7: 9926 ETA
8: 9926 ETA
9: 20514897 Basque_Country_(autonomous_community)
10: 404727 People's_Party_(Spain)
1: 20514897 Basque_Country_(autonomous_community)
 [pem_pos=2, final_scores=0.32902014,
final_local_scores=0.29961032, pem=-2.0219665,
similarity=24.026451]
2: 41188263 Madrid
 [pem_pos=0, final_scores=0.50515324,
final_local_scores=0.52284086, pem=-0.13236846,
similarity=20.412167]
3: 20514897 Basque_Country_(autonomous_community)
 [pem_pos=2, final_scores=0.34651083,
final_local_scores=0.3187163, pem=-2.0219665,
similarity=24.596611]
4: 9926 ETA
 [pem_pos=0, final_scores=0.5700515,
final_local_scores=0.54463625, pem=-0.10190937,
similarity=20.896944]
5: 539568 El_Mundo_(Spain)
 [pem_pos=0, final_scores=0.46915007,
final_local_scores=0.5366233, pem=0.0, similarity=20.103622]
6: 9926 ETA
 [pem_pos=0, final_scores=0.60977906,
final_local_scores=0.58803254, pem=-0.10190937,

```
similarity=22.191977]
7: 9926 ETA
    [pem_pos=0, final_scores=0.7283341,
final_local_scores=0.7175359, pem=0.0, similarity=25.502417]
8: 9926 ETA
    [pem_pos=0, final_scores=0.5676322,
final_local_scores=0.5419935, pem=-0.10190937,
similarity=20.818079]
9: 20514897 Basque_Country_(autonomous_community)
    [pem_pos=2, final_scores=0.31538635,
final_local_scores=0.2847175, pem=-2.0219665,
similarity=23.58202]
10: 404727 People's_Party_(Spain)
    [pem_pos=0, final_scores=0.3332047,
final_local_scores=0.32621482, pem=-0.36494392,
similarity=15.809236]
```

Sample: b'307_Thirty'

Thirty killed as floods plunge [1 1 Lahore 1] into chaos . [2 2 ISLAMABAD 2] 1996-08-24 At least 30 people have been killed and about 100 injured in the flood-hit [3 3 Pakistani 3] city of [4 4 Lahore 4] , newspapers reported on Saturday . They said 461 mm (18 inches) of rain had drenched the [5 5 Punjab 5] provincial capital in 36 hours , turning streets into rivers , knocking out power , water and telephone services , disrupting air and rail traffic , and sweeping away houses and cars . Newspapers quoted witnesses as saying they had seen bodies floating in the streets . Among the dead were five members of the religious [6 6 Jamaat-i-Islami 6] party who drowned while trying to remove books from a basement library . They said thousands of people had been made homeless after a breach opened in the city canal , inundating residential areas . Army troops were called in to evacuate residents of low-lying areas to higher ground . Officials said the [7 7 Ravi 7] and [8 8 Chenab 8] rivers , which both flow through [9 9 Punjab 9] , were in high flood and emergency services backed by troops were on full alert .

```
1: 125315 Lahore
2: 51153 Islamabad
3: 23235 Pakistan
4: 125315 Lahore
5: 24751 Punjab,_Pakistan
6: 676299 Jamaat-e-Islami_Pakistan
7: 815456 Ravi_River
8: 1100160 Chenab_River
9: 24751 Punjab,_Pakistan
1: 125315 Lahore
    [pem_pos=1, final_scores=0.41167998,
final_local_scores=0.3528163, pem=0.0, similarity=14.618456]
2: 51153 Islamabad
    [pem_pos=0, final_scores=0.40431935,
final_local_scores=0.3664332, pem=0.0, similarity=15.02481]
```

```
3: 23235 Pakistan
    [pem_pos=0, final_scores=0.5571191,
final_local_scores=0.5363576, pem=-0.49826378,
similarity=22.80533]
4: 125315 Lahore
    [pem_pos=0, final_scores=0.49019462,
final_local_scores=0.43858165, pem=0.0, similarity=17.177864]
5: 24751 Punjab, Pakistan
    [pem_pos=1, final_scores=0.3164528,
final_local_scores=0.24291193, pem=-1.088274,
similarity=17.256893]
6: 676299 Jamaat-e-Islami_Pakistan
    [pem_pos=0, final_scores=0.34385204,
final_local_scores=0.37478697, pem=-0.2876821,
similarity=16.838564]
7: 815456 Ravi_River
    [pem_pos=0, final_scores=0.5628479,
final_local_scores=0.59752023, pem=-0.4142058,
similarity=24.173424]
8: 1100160 Chenab_River
    [pem_pos=0, final_scores=0.4374138,
final_local_scores=0.45878273, pem=0.0, similarity=17.780704]
9: 24751 Punjab, Pakistan
    [pem_pos=1, final_scores=0.30855137,
final_local_scores=0.23428082, pem=-1.088274,
similarity=16.999323]
```

Sample: b'308_Internet'

Internet Startup funded to develop [1 1 Java 1] software . [2 2 MOUNTAIN VIEW 2] , [3 3 Calif. 3] 1996-08-25 A small team of engineers from [4 4 Sun Microsystems Inc. 4] 's JavaSoft unit said Sunday they have formed a new company , dubbed Internet Startup , to build [5 5 Java 5] infrastructure software . The fledgling company , established in a ground-floor office here over the last two weeks , has received venture financing from [6 6 Bessemer Venture Partners 6] of [7 7 Menlo Park 7] , [8 8 Calif 8] .. David Cowan , Internet Startup founder and acting chief executive , is a general partner of [9 9 Bessemer 9] . The startup company 's acting chairman is Jim Bidzos , the president of [10 10 RSA Data Security 10] , a unit of Security Dynamics Technologies Inc. as well as chairman of [11 11 VeriSign 11] . Internet Startup , which opened its doors with about half a dozen initial employees , combines experience at JavaSoft , 1 Apple Computer Inc. , and Oracle Systems . " [12 12 Java 12] portends dramatic changes in the way we use the [13 13 Internet 13] , " said Hong Bui , vice president of engineering of the new company after serving as a senior engineer at JavaSoft . [14 1 4 Java 14] is a computer programming language introduced by [1 5 15 Sun Microsystems 15] in mid-1995 which has immediately captured the attention of the industry for its ability to operate across virtually all computer system in a relatively

secure manner . Just last week , [16 16 Sun Microsystems 16] and the [17 17 Silicon Valley 17] venture capital giant [18 18 Kleiner Perkins Caufield & Byers 18] said they had completed financing of a \$ 100 million fund managed by [19 19 Kleiner Perkins 19] to fund startups developing [20 20 Java 20] technologies . [21 21 Java 21] has been licensed by nearly 50 organisations , ranging from [22 22 Microsoft Corp. 22] and [23 23 International Business Machines Corp. 23] to the [24 24 Taiwan 24] government . Prasad Wagle , another former senior JavaSoft engineer who is among the founding engineers at Internet Startup , said the new company aims to build software infrastructure using [25 25 Java 25] to make networked applications ubiquitous . One feature of the [26 26 Java 26] language is that small software programmes , known as " applets " because they are small applications , can be downloaded from the server computers at the centre of networks onto individual computers for use . In this model , individual computer users can always gain access to the latest programmes and do not need to store more software than they are currently using on their computers at any one time , also saving costs of memory and storage . Chris Zuleeg , a veteran of [27 27 Apple 27] and a former JavaSoft marketing manager , is vice president of marketing at Internet Startup , whose Web site is www.internetstartup.com . [28 28 Bessemer 28] has funded numerous [29 29 Internet 29] pioneers , including [30 30 PSI Net 30] , [31 31 VeriSign 31] and Individual .

```
1: 15881 Java_(programming_language)
2: 108161 Mountain_View,_California
3: 5407 California
4: 26980 Sun_Microsystems
5: 15881 Java_(programming_language)
6: 2828841 Bessemer_Venture_Partners
7: 105222 Menlo_Park,_California
8: 5407 California
9: 2828841 Bessemer_Venture_Partners
10: 301240 RSA_Security
11: 319632 Verisign
12: 15881 Java_(programming_language)
13: 33139 World_Wide_Web
14: 15881 Java_(programming_language)
15: 26980 Sun_Microsystems
16: 26980 Sun_Microsystems
17: 26976 Silicon_Valley
18: 299967 Kleiner_Perkins_Caufield_&_Byers
19: 299967 Kleiner_Perkins_Caufield_&_Byers
20: 15881 Java_(programming_language)
21: 15881 Java_(programming_language)
22: 19001 Microsoft
23: 40379651 IBM
24: 25734 Taiwan
25: 15881 Java_(programming_language)
```

```
26: 15881 Java_(programming_language)
27: 856 Apple_Inc.
28: 2828841 Bessemer_Venture_Partners
29: 33139 World_Wide_Web
30: 2661979 PSINet
31: 319632 Verisign
1: 856 Apple_Inc.
    [winner: 345676 Apple_Computer,_Inc._v._Microsoft_Corp.,
pem_pos=1, final_scores=-0.7374246,
final_local_scores=-0.74061984, pem=-0.6931472,
similarity=-14.242417]
    gt: 856 Apple_Inc., pem_pos=0, final_scores=-0.9844683,
final_local_scores=-1.1522844, pem=0.0,
similarity=-30.29675 ]
    [winner: 2531827 Inc._(magazine), pem_pos=0,
final_scores=-0.7808501, final_local_scores=-0.8330256,
pem=0.0, similarity=-20.769432]
    RECALL MISS: 856 Apple_Inc.]
1: 15881 Java_(programming_language)
    [pem_pos=0, final_scores=0.43349153,
final_local_scores=0.41978776, pem=-0.5985113,
similarity=19.871815]
2: 108161 Mountain_View,_California
    [pem_pos=0, final_scores=0.46009493,
final_local_scores=0.4711452, pem=-0.541827,
similarity=21.096165]
3: 5407 California
    [pem_pos=0, final_scores=0.48728043,
final_local_scores=0.5718795, pem=0.0, similarity=21.155739]
4: 26980 Sun_Microsystems
    [pem_pos=0, final_scores=0.48978782,
final_local_scores=0.44184452, pem=0.0, similarity=17.275234]
5: 15881 Java_(programming_language)
    [pem_pos=0, final_scores=0.42853063,
final_local_scores=0.41436875, pem=-0.5985113,
similarity=19.7101]
6: 2828841 Bessemer_Venture_Partners
    [pem_pos=0, final_scores=0.5625811,
final_local_scores=0.6650174, pem=0.0, similarity=23.93516]
7: 105222 Menlo_Park,_California
    [pem_pos=0, final_scores=0.35575444,
final_local_scores=0.411027, pem=0.0, similarity=16.35558]
8: 5407 California
    [pem_pos=0, final_scores=0.58451706,
final_local_scores=0.6780957, pem=0.0, similarity=24.325441]
9: 2828841 Bessemer_Venture_Partners
    [pem_pos=8, final_scores=0.52694577,
final_local_scores=0.6260912, pem=-2.187547,
similarity=34.669743]
10: 301240 RSA_Security
    [pem_pos=0, final_scores=0.37949347,
final_local_scores=0.4062221, pem=0.0, similarity=16.21219]
```

11: 319632 Verisign
[pem_pos=0, final_scores=0.50879866,
final_local_scores=0.5356885, pem=0.0, similarity=20.075727]
12: 15881 Java_(programming_language)
[pem_pos=0, final_scores=0.4304822,
final_local_scores=0.41650057, pem=-0.5985113,
similarity=19.773718]
13: 33139 World_Wide_Web
[pem_pos=1, final_scores=0.38444227,
final_local_scores=0.36128896, pem=-3.3911073,
similarity=33.312664]
14: 15881 Java_(programming_language)
[pem_pos=0, final_scores=0.458062,
final_local_scores=0.44662726, pem=-0.5985113,
similarity=20.672758]
15: 26980 Sun_Microsystems
[pem_pos=0, final_scores=0.60902643,
final_local_scores=0.5720946, pem=0.0, similarity=21.162157]
16: 26980 Sun_Microsystems
[pem_pos=0, final_scores=0.46605283,
final_local_scores=0.4159177, pem=0.0, similarity=16.501526]
17: 26976 Silicon_Valley
[pem_pos=0, final_scores=0.3598265,
final_local_scores=0.35344788, pem=0.0, similarity=14.637302]
18: 299967 Kleiner_Perkins_Caufield_&_Byers
[pem_pos=0, final_scores=0.44927114,
final_local_scores=0.48069584, pem=0.0, similarity=18.434635]
19: 299967 Kleiner_Perkins_Caufield_&_Byers
[pem_pos=0, final_scores=0.44257796,
final_local_scores=0.47338456, pem=0.0, similarity=18.216452]
20: 15881 Java_(programming_language)
[pem_pos=0, final_scores=0.46126294,
final_local_scores=0.45012385, pem=-0.5985113,
similarity=20.777103]
21: 15881 Java_(programming_language)
[pem_pos=0, final_scores=0.52181894,
final_local_scores=0.5162721, pem=-0.5985113,
similarity=22.7511]
22: 19001 Microsoft
[pem_pos=0, final_scores=0.5060273,
final_local_scores=0.46492016, pem=-0.006739087,
similarity=18.000507]
23: 40379651 IBM
[pem_pos=0, final_scores=0.39504915,
final_local_scores=0.35807395, pem=0.0, similarity=14.775353]
24: 25734 Taiwan
[pem_pos=0, final_scores=0.36512798,
final_local_scores=0.46737736, pem=-0.077886656,
similarity=18.460745]
25: 15881 Java_(programming_language)
[pem_pos=0, final_scores=0.4132653,
final_local_scores=0.39769375, pem=-0.5985113,

```
similarity=19.212484]
26: 15881 Java_(programming_language)
    [pem_pos=0, final_scores=0.43073684,
final_local_scores=0.41677868, pem=-0.5985113,
similarity=19.782017]
27: 856 Apple_Inc.
    [pem_pos=0, final_scores=0.43677837,
final_local_scores=0.40447766, pem=-0.40360686,
similarity=18.35501]
28: 2828841 Bessemer_Venture_Partners
    [pem_pos=8, final_scores=0.5716542,
final_local_scores=0.67492837, pem=-2.187547,
similarity=36.127144]
29: 33139 World_Wide_Web
    [pem_pos=1, final_scores=0.38054776,
final_local_scores=0.35703486, pem=-3.3911073,
similarity=33.185715]
30: 2661979 PSINet
    [pem_pos=0, final_scores=0.41243845,
final_local_scores=0.4777695, pem=0.0, similarity=18.347307]
31: 319632 Verisign
    [pem_pos=0, final_scores=0.6259301,
final_local_scores=0.6636368, pem=0.0, similarity=23.893959]
```

Sample: b'309_GOLF'

GOLF - [1 1 MICKELSON 1] WINS FOURTH TITLE OF YEAR IN [2 2 AK RON 2] . [3 3 AKRON 3] , [4 4 Ohio 4] 1996-08-25 [5 5 Phil Mickelson 5] birdied two of the last three holes to win [6 6 World Series of Golf 6] by three strokes over [7 7 Billy Mayfair 7] on Sunday . It was the fourth tournament title this year for [8 8 Mickelson 8] , who shot an even-par 70 , after being tied for the lead with [9 9 Billy Mayfair 9] with three holes to play . Along with Mayfiar at 277 for the tournament were [10 10 Steve Stricker 10] , who had a 68 , and [11 11 Duffy Waldorf 11] , with a 66 . " It was very hard to sleep last night because there was so much I could accomplish with this win , " said [12 12 Mickelson 12] , who had a three-stroke lead entering the third round . " This was a win I wanted very , very much . " [13 13 Mickelson 13] 's victory gave him a 10 year exemption to the [14 14 PGA Tour 14] . The \$ 378,000 first place check brings [15 15 Mickelson 15] back to the top of the money list with \$ 1,574,799 won this year . " This is a major championship golf course , and for me to perform well on this style of course is a big step up for me in my career and my performance in future majors , " he said . [16 16 Mickelson 16] three-stroke lead was cut to two when [17 17 Mayfair 17] birdied the course 's only easy hole , the par five second hole , while [18 18 Mickelson 18] three-putted for par from 25 feet . On the back nine [19 19 Mickelson 19] began driving erratically , and poor tee shots resulted in bogeys on the , eighth , 12th and 13th holes , bringing [20 20 Mickelson 20] back to four under par , tied with [21 21 Mayfair 21]

, who had parred 14 straight holes after the birdie on no.2 . [22 22 Mickelson 22] then set up a tap in birdie on the 16th , sending a wedge shot to 18 inches . He had another birdie on the 17th , where he his a 6-iron to six feet . [23 23 Mayfair 23] bogeyed the 17th , missing a five foot par putt , and dropped from solo second place to a three way tie for second . It was the second successive year in which [24 24 Mayfair 24] has finished runner up in this tournament . He lost to [25 25 Greg Norman 25] in sudden death last year . The defending champion was in contention , two behind [26 26 Mickelson 26] for much of the day , until he bogeyed the 13th and 14th holes .

```
1: 592392 Phil_Mickelson
2: 129842 Akron,_Ohio
3: 129842 Akron,_Ohio
4: 22199 Ohio
5: 592392 Phil_Mickelson
6: 1536678 World_Series_of_Golf
7: 2080225 Billy_Mayfair
8: 592392 Phil_Mickelson
9: 2080225 Billy_Mayfair
10: 2310551 Steve_Stricker
11: 1781343 Duffy_Waldorf
12: 592392 Phil_Mickelson
13: 592392 Phil_Mickelson
14: 23546455 PGA_Tour
15: 592392 Phil_Mickelson
16: 592392 Phil_Mickelson
17: 2080225 Billy_Mayfair
18: 592392 Phil_Mickelson
19: 592392 Phil_Mickelson
20: 592392 Phil_Mickelson
21: 2080225 Billy_Mayfair
22: 592392 Phil_Mickelson
23: 2080225 Billy_Mayfair
24: 2080225 Billy_Mayfair
25: 157447 Greg_Norman
26: 592392 Phil_Mickelson
1: 592392 Phil_Mickelson
    [pem_pos=0, final_scores=0.6644639,
final_local_scores=0.6152025, pem=0.0, similarity=22.448582]
2: 129842 Akron,_Ohio
    [pem_pos=0, final_scores=0.34359962,
final_local_scores=0.4526406, pem=-0.6020413,
similarity=20.871407]
3: 129842 Akron,_Ohio
    [pem_pos=0, final_scores=0.51660544,
final_local_scores=0.64162314, pem=-0.6020413,
similarity=26.511023]
4: 22199 Ohio
    [pem_pos=0, final_scores=0.5156578,
final_local_scores=0.6825539, pem=-0.19147073,
```

```
similarity=25.49973]
5: 592392 Phil_Mickelson
    [pem_pos=0, final_scores=0.8609367,
final_local_scores=0.82884663, pem=0.0, similarity=28.82415]
6: 1536678 World_Series_of_Golf
    [pem_pos=0, final_scores=0.35553312,
final_local_scores=0.43107158, pem=0.0, similarity=16.953749]
7: 2080225 Billy_Mayfair
    [pem_pos=0, final_scores=0.8254897,
final_local_scores=0.8418862, pem=0.0, similarity=29.213278]
8: 592392 Phil_Mickelson
    [pem_pos=0, final_scores=0.7204656,
final_local_scores=0.67637587, pem=0.0, similarity=24.274118]
9: 2080225 Billy_Mayfair
    [pem_pos=0, final_scores=0.83131224,
final_local_scores=0.8482464, pem=0.0, similarity=29.403078]
10: 2310551 Steve_Stricker
    [pem_pos=0, final_scores=0.7478482,
final_local_scores=0.76183987, pem=0.0, similarity=26.824533]
11: 1781343 Duffy_Waldorf
    [pem_pos=0, final_scores=0.8022986,
final_local_scores=0.89650196, pem=0.0, similarity=30.84312]
12: 592392 Phil_Mickelson
    [pem_pos=0, final_scores=0.8382384,
final_local_scores=0.8050248, pem=0.0, similarity=28.113258]
13: 592392 Phil_Mickelson
    [pem_pos=0, final_scores=0.840807,
final_local_scores=0.8078305, pem=0.0, similarity=28.196987]
14: 23546455 PGA_Tour
    [pem_pos=0, final_scores=0.36034268,
final_local_scores=0.3437732, pem=0.0, similarity=14.348591]
15: 592392 Phil_Mickelson
    [pem_pos=0, final_scores=0.86779445,
final_local_scores=0.83731025, pem=0.0, similarity=29.076721]
16: 592392 Phil_Mickelson
    [pem_pos=0, final_scores=0.75589377,
final_local_scores=0.7150757, pem=0.0, similarity=25.428997]
17: 2080225 Billy_Mayfair
    [pem_pos=0, final_scores=1.2732766,
final_local_scores=1.3347483, pem=0.0, similarity=43.921265]
18: 592392 Phil_Mickelson
    [pem_pos=0, final_scores=0.72674847,
final_local_scores=0.6832389, pem=0.0, similarity=24.478926]
19: 592392 Phil_Mickelson
    [pem_pos=0, final_scores=0.72231025,
final_local_scores=0.67839086, pem=0.0, similarity=24.33425]
20: 592392 Phil_Mickelson
    [pem_pos=0, final_scores=0.7195779,
final_local_scores=0.6754062, pem=0.0, similarity=24.245182]
21: 2080225 Billy_Mayfair
    [pem_pos=0, final_scores=1.0043443,
final_local_scores=1.0409808, pem=0.0, similarity=35.154655]
```

```
22: 592392 Phil_Mickelson
    [pem_pos=0, final_scores=0.80977017,
final_local_scores=0.77392757, pem=0.0, similarity=27.185255]
23: 2080225 Billy_Mayfair
    [pem_pos=0, final_scores=1.2272834,
final_local_scores=1.2845078, pem=0.0, similarity=42.42199]
24: 2080225 Billy_Mayfair
    [pem_pos=0, final_scores=1.2498482,
final_local_scores=1.3091564, pem=0.0, similarity=43.15755]
25: 157447 Greg_Norman
    [pem_pos=0, final_scores=0.6121537,
final_local_scores=0.6161681, pem=0.0, similarity=22.477398]
26: 592392 Phil_Mickelson
    [pem_pos=0, final_scores=0.82605535,
final_local_scores=0.79171664, pem=0.0, similarity=27.716116]
```

Sample: b'310_GOLF'

GOLF - SCORES AT THE [1 1 WORLD SERIES OF GOLF 1] . [2 2 AKR
ON 2] , [3 3 Ohio 3] 1996-08-24 Scores after the final round
of the \$ 2.1 million [4 4 NEC World Series of Golf 4] at
Firestone C.C , 7149 yards , par 70 (players [5 5 U.S. 5]
unless noted) : 274 [6 6 Phil Mickelson 6] 70 66 68 70 277 [7
7 Duffy Waldorf 7] 70 70 71 66 , [8 8 Steve Stricker 8] 68 72
69 68 , [9 9 Billy Mayfair 9] 66 71 70 70 278 [10 10 Greg Nor
man 10] ([11 11 Australia 11]) 70 68 69 71 280 [12 12 Alexa
nder Cejka 12] ([13 13 Germany 13]) 72 71 71 66 , [14 14 Da
vis Love 14] 70 74 67 69 281 [15 15 John Cook 15] 70 69 71 71
282 [16 16 Corey Pavin 16] 73 70 70 69 283 [17 17 Tom Lehman 1
7] 72 69 74 68 , [18 18 Fred Funk 18] 72 70 73 68 , [19 19 Ma
rk Brooks 19] 69 69 74 71 , [20 20 Nick Faldo 20] ([21 21 Br
itain 21]) 70 71 68 74 284 [22 22 D.A. Weibring 22] 73 69 74
68 , [23 23 Tim Herron 23] 70 67 75 72 , Mark O'Meara 73 71
69 71 , [24 24 Jim Furyk 24] 75 69 67 73 , [25 25 Justin Leon
ard 25] 69 70 71 74 285 [26 26 Loren Roberts 26] 72 73 71
69 , [27 27 Hal Sutton 27] 72 69 74 70 , [28 28 Fred Couples 2
8] 73 68 72 72 , [29 29 Craig Stadler 29] 73 72 67 73 286 [30
30 Hidemichi Tanaka 30] ([31 31 Japan 31]) 66 75 75 70 , [32
32 Steve Jones 32] 70 69 76 71 , [33 33 Paul Goydos 33] 66 75
74 71 , [34 34 Ernie Els 34] ([35 35 South Africa 35]) 71
71 71 73 287 [36 36 Costantino Rocca 36] ([37 37 Italy 37])
74 71 75 67 , [38 38 Clarence Rose 38] 72 71 72 72 , [39 39 Cr
raig Parry 39] ([40 40 Australia 40]) 73 75 67 72 , [41 41 Wi
llie Wood 41] 75 69 69 74 288 [42 42 Shigeki Maruyama 42] ([4
3 43 Japan 43]) 75 71 70 72 , [44 44 Anders Forsbrand 44] ([
45 45 Sweden 45]) 70 75 71 72 289 [46 46 Scott Hoch 46] 71
68 77 73 290 [47 47 Tom Watson 47] 79 70 68 73 292 [48 48 Way
ne Westner 48] ([49 49 South Africa 49]) 77 68 73 74 , [50 5
0 Sven Struver 50] ([51 51 Germany 51]) 72 72 72 76 294 [52
52 Satoshi Higashi 52] ([53 53 Japan 53]) 75 72 74 73 , [54
54 Scott McCarron 54] 76 70 74 74 295 [55 55 Stewart Ginn 55]
([56 56 Australia 56]) 73 72 77 73 298 Steve Schneiter 77
74 76 71 , [57 57 Paul Stankowski 57] 74 75 74 75 , Seiki

Okuda ([58 58 Japan 58]) 81 70 72 75 301 [59 59 Brad Bryant
59] 73 72 77 79
1: 1536678 World_Series_of_Golf
2: 129842 Akron,_Ohio
3: 22199 Ohio
4: 1527695 WGC-Bridgestone_Invitational
5: 3434750 United_States
6: 592392 Phil_Mickelson
7: 1781343 Duffy_Waldorf
8: 2310551 Steve_Stricker
9: 2080225 Billy_Mayfair
10: 157447 Greg_Norman
11: 4689264 Australia
12: 2080183 Alex_Cejka
13: 11867 Germany
14: 1388818 Davis_Love_III
15: 2810777 John_Cook_(golfer)
16: 1530024 Corey_Pavin
17: 1516564 Tom_Lehman
18: 1646050 Fred_Funk
19: 2295891 Mark_Brooks_(golfer)
20: 759285 Nick_Faldo
21: 31717 United_Kingdom
22: 5653062 D._A._Weibring
23: 5242796 Tim_Herron
24: 247955 Jim_Furyk
25: 1507170 Justin_Leonard
26: 2567564 Loren_Roberts
27: 654136 Hal_Sutton
28: 1221065 Fred_Couples
29: 1261748 Craig_Stadler
30: 2114665 Hidemichi_Tanaka
31: 15573 Japan
32: 1924538 Steve_Jones_(golfer)
33: 6280981 Paul_Goydos
34: 230467 Ernie_Els
35: 17416221 South_Africa
36: 2157151 Costantino_Rocca
37: 14532 Italy
38: 8385080 Clarence_Rose
39: 3738086 Craig_Parry
40: 4689264 Australia
41: 7018258 Willie_Wood_(golfer)
42: 414220 Shigeki_Maruyama
43: 15573 Japan
44: 4377636 Anders_Forsbrand
45: 5058739 Sweden
46: 3396339 Scott_Hoch
47: 585794 Tom_Watson_(golfer)
48: 5206729 Wayne_Westner
49: 17416221 South_Africa
50: 4337089 Sven_Strüver

```
51: 11867 Germany
52: 21138421 Satoshi_Higashi
53: 15573 Japan
54: 5845032 Scott_McCarron
55: 5418317 Stewart_Ginn
56: 4689264 Australia
57: 6189303 Paul_Stankowski
58: 15573 Japan
59: 4450518 Brad_Bryant
1: 1536678 World_Series_of_Golf
    [pem_pos=0, final_scores=0.41443664,
final_local_scores=0.48188168, pem=0.0, similarity=18.470024]
2: 129842 Akron,_Ohio
    [pem_pos=0, final_scores=0.563028,
final_local_scores=0.6313571, pem=-0.6020413,
similarity=26.204662]
3: 22199 Ohio
    [pem_pos=0, final_scores=0.5776347,
final_local_scores=0.66761744, pem=-0.19147073,
similarity=25.053997]
4: 1527695 WGC-Bridgestone_Invitational
    [pem_pos=0, final_scores=0.39023727,
final_local_scores=0.4056993, pem=0.0, similarity=16.19659]
5: 3434750 United_States
    [pem_pos=0, final_scores=0.45395464,
final_local_scores=0.50860935, pem=-0.14075686,
similarity=20.033087]
6: 592392 Phil_Mickelson
    [pem_pos=0, final_scores=0.73412675,
final_local_scores=0.75448394, pem=0.0, similarity=26.605019]
7: 1781343 Duffy_Waldorf
    [pem_pos=0, final_scores=0.6659506,
final_local_scores=0.74649477, pem=0.0, similarity=26.366606]
8: 2310551 Steve_Stricker
    [pem_pos=0, final_scores=0.6528185,
final_local_scores=0.67852074, pem=0.0, similarity=24.338125]
9: 2080225 Billy_Mayfair
    [pem_pos=0, final_scores=0.7761298,
final_local_scores=0.82621825, pem=0.0, similarity=28.745714]
10: 157447 Greg_Norman
    [pem_pos=0, final_scores=0.6208721,
final_local_scores=0.63135666, pem=0.0, similarity=22.930655]
11: 4689264 Australia
    [pem_pos=0, final_scores=0.5903614,
final_local_scores=0.662932, pem=-0.17168477,
similarity=24.806574]
12: 2080183 Alex_Čejka
    [pem_pos=0, final_scores=0.7696019,
final_local_scores=0.8254357, pem=0.0, similarity=28.72236]
13: 11867 Germany
    [pem_pos=0, final_scores=0.8014119,
final_local_scores=0.9068156, pem=-0.15529114,
```

```
similarity=31.995396]
14: 1388818 Davis_Love_III
    [pem_pos=0, final_scores=0.63649935,
final_local_scores=0.64831096, pem=0.0, similarity=23.436605]
15: 2810777 John_Cook_(golfer)
    [pem_pos=1, final_scores=0.5664092,
final_local_scores=0.5773029, pem=-1.1074443,
similarity=27.340036]
16: 1530024 Corey_Pavin
    [pem_pos=0, final_scores=0.87961704,
final_local_scores=0.91253346, pem=0.0, similarity=31.321533]
17: 1516564 Tom_Lehman
    [pem_pos=0, final_scores=0.88229114,
final_local_scores=0.91437024, pem=0.0, similarity=31.376345]
18: 1646050 Fred_Funk
    [pem_pos=0, final_scores=0.6054161,
final_local_scores=0.63334197, pem=0.0, similarity=22.9899]
19: 2295891 Mark_Brooks_(golfer)
    [pem_pos=0, final_scores=0.616003,
final_local_scores=0.65722704, pem=-0.17831625,
similarity=24.67239]
20: 759285 Nick_Faldo
    [pem_pos=0, final_scores=0.8622995,
final_local_scores=0.89302415, pem=0.0, similarity=30.739336]
21: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.50400066,
final_local_scores=0.58613473, pem=-0.61114657,
similarity=24.904652]
22: 5653062 D._A._Weibring
    [pem_pos=0, final_scores=0.6132972,
final_local_scores=0.66822135, pem=0.0, similarity=24.030771]
23: 5242796 Tim_Herron
    [pem_pos=0, final_scores=0.7073932,
final_local_scores=0.75978184, pem=0.0, similarity=26.763119]
24: 247955 Jim_Furyk
    [pem_pos=0, final_scores=0.7922798,
final_local_scores=0.8192631, pem=0.0, similarity=28.538158]
25: 1507170 Justin_Leonard
    [pem_pos=0, final_scores=0.7105084,
final_local_scores=0.742996, pem=0.0, similarity=26.262196]
26: 2567564 Loren_Roberts
    [pem_pos=0, final_scores=0.6066273,
final_local_scores=0.6282886, pem=0.0, similarity=22.839098]
27: 654136 Hal_Sutton
    [pem_pos=0, final_scores=0.623342,
final_local_scores=0.6555541, pem=0.0, similarity=23.652756]
28: 1221065 Fred_Couples
    [pem_pos=0, final_scores=0.7327621,
final_local_scores=0.76007456, pem=0.0, similarity=26.771854]
29: 1261748 Craig_Stadler
    [pem_pos=0, final_scores=0.658457,
final_local_scores=0.6879941, pem=0.0, similarity=24.62083]
```

30: 2114665 Hidemichi_Tanaka
[pem_pos=0, final_scores=0.48021132,
final_local_scores=0.61522645, pem=0.0, similarity=22.449297]
31: 15573 Japan
[pem_pos=0, final_scores=0.7086404,
final_local_scores=0.79716456, pem=-0.13549899,
similarity=28.615557]
32: 1924538 Steve_Jones_(golfer)
[pem_pos=4, final_scores=0.40541822,
final_local_scores=0.4248876, pem=-2.187635,
similarity=28.665905]
33: 6280981 Paul_Goydos
[pem_pos=0, final_scores=0.7456257,
final_local_scores=0.7931821, pem=0.0, similarity=27.759848]
34: 230467 Ernie_Els
[pem_pos=0, final_scores=0.7142089,
final_local_scores=0.7427444, pem=0.0, similarity=26.254686]
35: 17416221 South_Africa
[pem_pos=0, final_scores=0.54083204,
final_local_scores=0.6246612, pem=-0.26251382,
similarity=24.15844]
36: 2157151 Costantino_Rocca
[pem_pos=0, final_scores=0.8109148,
final_local_scores=0.86915, pem=0.0, similarity=30.026882]
37: 14532 Italy
[pem_pos=0, final_scores=0.6625368,
final_local_scores=0.77674747, pem=-0.20612453,
similarity=28.390345]
38: 8385080 Clarence_Rose
[pem_pos=0, final_scores=0.4867081,
final_local_scores=0.5708775, pem=0.0, similarity=21.125835]
39: 3738086 Craig_Parry
[pem_pos=0, final_scores=0.73012954,
final_local_scores=0.764519, pem=0.0, similarity=26.904484]
40: 4689264 Australia
[pem_pos=0, final_scores=0.6369472,
final_local_scores=0.71381986, pem=-0.17168477,
similarity=26.32517]
41: 7018258 Willie_Wood_(golfer)
[pem_pos=1, final_scores=0.592948,
final_local_scores=0.6399195, pem=-0.86820155,
similarity=27.907602]
42: 414220 Shigeki_Maruyama
[pem_pos=0, final_scores=0.58330697,
final_local_scores=0.6283722, pem=0.0, similarity=22.841593]
43: 15573 Japan
[pem_pos=0, final_scores=0.69933814,
final_local_scores=0.7870032, pem=-0.13549899,
similarity=28.312323]
44: 4377636 Anders_Forsbrand
[pem_pos=0, final_scores=0.53688735,
final_local_scores=0.59165263, pem=0.0, similarity=21.745808]

45: 5058739 Sweden
[pem_pos=0, final_scores=0.53013587,
final_local_scores=0.59807706, pem=-0.2135508,
similarity=23.098846]
46: 3396339 Scott_Hoch
[pem_pos=0, final_scores=0.70375806,
final_local_scores=0.7529266, pem=0.0, similarity=26.558544]
47: 585794 Tom_Watson_(golfer)
[pem_pos=0, final_scores=0.6465782,
final_local_scores=0.66344947, pem=-0.48845047,
similarity=26.544638]
48: 5206729 Wayne_Westner
[pem_pos=0, final_scores=0.7078373,
final_local_scores=0.79269034, pem=0.0, similarity=27.745174]
49: 17416221 South_Africa
[pem_pos=0, final_scores=0.5470653,
final_local_scores=0.63147014, pem=-0.26251382,
similarity=24.361633]
50: 4337089 Sven_Strüver
[pem_pos=0, final_scores=0.69521123,
final_local_scores=0.79461974, pem=0.0, similarity=27.802752]
51: 11867 Germany
[pem_pos=0, final_scores=0.6705359,
final_local_scores=0.76385343, pem=-0.15529114,
similarity=27.72912]
52: 21138421 Satoshi_Higashi
[pem_pos=0, final_scores=0.6017222,
final_local_scores=0.71844494, pem=0.0, similarity=25.529543]
53: 15573 Japan
[pem_pos=0, final_scores=0.73590904,
final_local_scores=0.82695144, pem=-0.13549899,
similarity=29.504457]
54: 5845032 Scott_McCarron
[pem_pos=0, final_scores=0.70153946,
final_local_scores=0.76913196, pem=0.0, similarity=27.042145]
55: 5418317 Stewart_Ginn
[pem_pos=0, final_scores=0.42206573,
final_local_scores=0.5175367, pem=0.0, similarity=19.53404]
56: 4689264 Australia
[pem_pos=0, final_scores=0.64554197,
final_local_scores=0.72320825, pem=-0.17168477,
similarity=26.60534]
57: 6189303 Paul_Stankowski
[pem_pos=0, final_scores=0.63552094,
final_local_scores=0.71657556, pem=0.0, similarity=25.473757]
58: 15573 Japan
[pem_pos=0, final_scores=0.73399025,
final_local_scores=0.8248554, pem=-0.13549899,
similarity=29.441906]
59: 4450518 Brad_Bryant
[pem_pos=0, final_scores=0.604188,
final_local_scores=0.645247, pem=0.0, similarity=23.34517]

Sample: b'311_TENNIS'

TENNIS - RESULTS AT HAMLET CUP . [1 1 COMMACK 1] , [2 2 New York 2] 1996-08-24 Results at the Hamlet Cup tennis tournament on Sunday (prefix number denotes seedings : Finals , singles 5 - [3 3 Andrei Medvedev 3] ([4 4 Ukraine 4]) beat [5 5 Martin Damm 5] ([6 6 Czech Republic 6]) 7-5 6-3 Finals , doubles [7 7 Luke Jensen 7] and [8 8 Murphy Jensen 8] ([9 9 U.S. 9]) beat [10 10 Alexander Volkov 10] ([11 11 Russia 11]) and Handrik Dreekmann ([12 12 Germany 12]) 6-3 7-6 (7-5)

1: 127266 Commack,_New_York
2: 645042 New_York_City
3: 702849 Andriy_Medvedev
4: 31750 Ukraine
5: 4026614 Martin_Damm
6: 5321 Czech_Republic
7: 6554040 Luke_Jensen
8: 6371949 Murphy_Jensen
9: 3434750 United_States
10: 1856445 Alexander_Vladimirovich_Volkov
11: 25391 Russia
12: 11867 Germany

1: 127266 Commack,_New_York
[pem_pos=0, final_scores=0.47514564,
final_local_scores=0.6026813, pem=0.0, similarity=22.074924]
2: 645042 New_York_City
[pem_pos=1, final_scores=0.20454578,
final_local_scores=0.22095972, pem=-1.7906687,
similarity=20.421526]
3: 702849 Andriy_Medvedev
[pem_pos=0, final_scores=0.8622014,
final_local_scores=0.98117703, pem=0.0, similarity=33.36999]
4: 31750 Ukraine
[pem_pos=0, final_scores=0.6030423,
final_local_scores=0.6364546, pem=-0.35407752,
similarity=25.008316]
5: 4026614 Martin_Damm
[pem_pos=0, final_scores=0.67170596,
final_local_scores=0.7931926, pem=0.0, similarity=27.760162]
6: 5321 Czech_Republic
[pem_pos=0, final_scores=0.5668322,
final_local_scores=0.5798492, pem=-0.27794555,
similarity=22.90508]
7: 6554040 Luke_Jensen
[pem_pos=0, final_scores=0.56821644,
final_local_scores=0.71339715, pem=0.0, similarity=25.378906]
8: 6371949 Murphy_Jensen
[pem_pos=0, final_scores=0.5682749,
final_local_scores=0.6749143, pem=0.0, similarity=24.230501]
9: 3434750 United_States
[pem_pos=0, final_scores=0.5037206,
final_local_scores=0.51193804, pem=-0.14075686,

```
similarity=20.132421]
10: 1856445 Alexander_Vladimirovich_Volkov
    [pem_pos=0, final_scores=0.4781838,
final_local_scores=0.5898029, pem=-0.6501985,
similarity=25.22649]
11: 25391 Russia
    [pem_pos=0, final_scores=0.7494708,
final_local_scores=0.80962867, pem=-0.20299491,
similarity=29.354565]
12: 11867 Germany
    [pem_pos=0, final_scores=0.88979775,
final_local_scores=0.9309277, pem=-0.15529114,
similarity=32.71495]

Sample: b'312_TENNIS'
TENNIS - RESULTS AT [1 1 TOSHIBA CLASSIC 1] . [2 2 CARLSBAD 2
] , [3 3 California 3] 1996-08-25 Results from the $ 450,000 [
4 4 Toshiba Classic 4] tennis tournament on Sunday ( prefix
number denotes seeding ) : Finals : 4 - [5 5 Kimiko Date 5] (
[6 6 Japan 6] ) beat 1 - [7 7 Arantxa Sanchez Vicario 7] ( [8
8 Spain 8] ) 3-6 6-3 6-0 .
1: 31726664 Southern_California_Open
2: 108030 Carlsbad,_California
3: 5407 California
4: 31726664 Southern_California_Open
5: 495296 Kimiko_Date-Krumm
6: 15573 Japan
7: 534624 Arantxa_Sánchez_Vicario
8: 26667 Spain
1: 31726664 Southern_California_Open
    [pem_pos=0, final_scores=0.47438794,
final_local_scores=0.4963218, pem=0.0, similarity=18.900946]
2: 108030 Carlsbad,_California
    [pem_pos=0, final_scores=0.520433,
final_local_scores=0.5680338, pem=-0.36353475,
similarity=23.017933]
3: 5407 California
    [pem_pos=0, final_scores=0.61535823,
final_local_scores=0.69595945, pem=-0.08662026,
similarity=25.329586]
4: 31726664 Southern_California_Open
    [pem_pos=0, final_scores=0.55578107,
final_local_scores=0.5852315, pem=0.0, similarity=21.554188]
5: 495296 Kimiko_Date-Krumm
    [pem_pos=0, final_scores=0.7374974,
final_local_scores=0.82487434, pem=0.0, similarity=28.705608]
6: 15573 Japan
    [pem_pos=0, final_scores=0.6437717,
final_local_scores=0.73024774, pem=-0.13549899,
similarity=26.618626]
7: 534624 Arantxa_Sánchez_Vicario
    [pem_pos=0, final_scores=0.5773526,
```

```
final_local_scores=0.65955806, pem=0.0, similarity=23.772242]
8: 26667 Spain
    [pem_pos=0, final_scores=0.5487438,
final_local_scores=0.6157821, pem=-0.18857051,
similarity=23.491354]

Sample: b'313_RALLYING'
RALLYING - LEADING POSITIONS IN 1,000 LAKES RALLY .
JYVASKLYA , [1 1 Finland 1] 1996-08-25 Leading positions on
Sunday after 23 special stages in the 1,000 Lakes Rally ,
sixth round of the world championship : 1. Tommi Makinen ( [2
2 Finland 2] ) [3 3 Mitsubishi Lancer 3] three hours eight
minutes one second 2. [4 4 Juha Kankkunen 4] ( [5 5 Finland 5]
) [6 6 Toyota Celica 6] 12 seconds behind 3. [7 7 Marcus Gron
holm 7] ( [8 8 Finland 8] ) [9 9 Toyota Celica 9] 2:09 4. [10
10 Jarmo Kytölehto 10] ( [11 11 Finland 11] ) Ford Escort
2:23 5. [12 12 Kenneth Eriksson 12] ( [13 13 Sweden 13] ) [14
14 Subaru Impreza 14] 2:39 6. [15 15 Carlos Sainz 15] ( [16 16
Spain 16] ) Ford Escort 3:03
1: 10577 Finland
2: 10577 Finland
3: 1119339 Mitsubishi_Lancer
4: 412742 Juha_Kankkunen
5: 10577 Finland
6: 172542 Toyota_Celica
7: 944600 Marcus_Grönholm
8: 10577 Finland
9: 172542 Toyota_Celica
10: 20730137 Jarmo_Kytölehto
11: 10577 Finland
12: 1411702 Kenneth_Eriksson
13: 5058739 Sweden
14: 527976 Subaru_Impreza
15: 407139 Carlos_Sainz
16: 26667 Spain
1: 10577 Finland
    [pem_pos=0, final_scores=0.8007141,
final_local_scores=0.82003796, pem=-0.2445894,
similarity=29.891396]
2: 10577 Finland
    [pem_pos=0, final_scores=0.7481046,
final_local_scores=0.76257014, pem=-0.2445894,
similarity=28.176443]
3: 1119339 Mitsubishi_Lancer
    [pem_pos=0, final_scores=0.29752254,
final_local_scores=0.3506856, pem=0.0, similarity=14.554871]
4: 412742 Juha_Kankkunen
    [pem_pos=0, final_scores=0.6842845,
final_local_scores=0.73944664, pem=0.0, similarity=26.156275]
5: 10577 Finland
    [pem_pos=0, final_scores=0.7705805,
final_local_scores=0.78712165, pem=-0.2445894,
```

```
similarity=28.90911]
6: 172542 Toyota_Celica
    [pem_pos=0, final_scores=0.3901319,
final_local_scores=0.41014582, pem=0.0, similarity=16.329283]
7: 944600 Marcus_Grönholm
    [pem_pos=0, final_scores=0.8232222,
final_local_scores=0.87351483, pem=0.0, similarity=30.157139]
8: 10577 Finland
    [pem_pos=0, final_scores=0.7673419,
final_local_scores=0.783584, pem=-0.2445894,
similarity=28.803537]
9: 172542 Toyota_Celica
    [pem_pos=0, final_scores=0.4555006,
final_local_scores=0.48155117, pem=0.0, similarity=18.460161]
10: 20730137 Jarmo_Kytölehto
    [pem_pos=0, final_scores=0.5241156,
final_local_scores=0.6577041, pem=0.0, similarity=23.716915]
11: 10577 Finland
    [pem_pos=0, final_scores=0.718376,
final_local_scores=0.73009616, pem=-0.2445894,
similarity=27.207354]
12: 1411702 Kenneth_Eriksson
    [pem_pos=0, final_scores=0.56908816,
final_local_scores=0.6397192, pem=0.0, similarity=23.18021]
13: 5058739 Sweden
    [pem_pos=0, final_scores=0.7586307,
final_local_scores=0.79708815, pem=-0.2135508,
similarity=29.037735]
14: 527976 Subaru_Impreza
    [pem_pos=0, final_scores=0.1714391,
final_local_scores=0.17285573, pem=0.0, similarity=9.248073]
15: 407139 Carlos_Sainz
    [pem_pos=0, final_scores=0.45289683,
final_local_scores=0.48857862, pem=0.0, similarity=18.669874]
16: 26667 Spain
    [pem_pos=0, final_scores=0.5857602,
final_local_scores=0.6489485, pem=-0.18857051,
similarity=24.481106]
```

Sample: b'314_MOTOCROSS'

MOTOCROSS - [1 1 SWEDISH 1] 500CC GRAND PRIX RESULTS . [2 2 L
ANDSKRONA 2] , [3 3 Sweden 3] 1996-08-25 Leading results in
the [4 4 Swedish 4] 500cc motocross Grand Prix on Sunday :
First race 1. [5 5 Joel Smets 5] ([6 6 Belgium 6]) [7 7 Hus
aberg 7] 2. [8 8 Peter Johansson 8] ([9 9 Sweden 9]) [10 10
Husqvarna 10] 3. Gert Jan Van Doorn ([11 11 Netherlands 11])
[12 12 Honda 12] 4. [13 13 Jacky Martens 13] ([14 14 Belgium
14]) [15 15 Husqvarna 15] 5. Peter Dirkx ([16 16 Belgium 16]
) [17 17 KTM 17] 6. Danny Theybers ([18 18 Belgium 18]) [19
19 Honda 19] Second race 1. [20 20 Shayne King 20] ([21 21 Ne
w Zealand 21]) [22 22 KTM 22] 2. [23 23 Martens 23] 3.
Theybers 4. Johan Boonen ([24 24 Belgium 24]) [25 25 Husqva

rna 25] 5. Dietmar Lalcher ([26 26 Germany 26]) [27 27 Honda
27] 6. Claus Manne Nielsen ([28 28 Denmark 28]) [29 29 KTM 2
9] Overall on day : 1. [30 30 Martens 30] 30 points 2. [31 31
Shayne King 31] 28 3. Smets 27 4. Theybers 25 5. Van Doorn 24
6. [32 32 Johansson 32] 17 World championship standings
(after 11 of 12 rounds) : 1. [33 33 Shayne King 33] 323
points 2. Smets 290 3. [34 34 Johansson 34] 236 4. Lacher 219
5. Darryll King ([35 35 New Zealand 35]) [36 36 Honda 36]
178 6. Van Doorn 176
1: 5058739 Sweden
2: 1019545 Landskrona
3: 5058739 Sweden
4: 5058739 Sweden
5: 6369246 Joël_Smets
6: 3343 Belgium
7: 3347463 Husaberg
8: 18781780 Peter_Johansson_(motorcyclist)
9: 5058739 Sweden
10: 13504069 Husqvarna_Motorcycles
11: 21148 Netherlands
12: 13729 Honda
13: 14948713 Jacky_Martens
14: 3343 Belgium
15: 13504069 Husqvarna_Motorcycles
16: 3343 Belgium
17: 243320 KTM
18: 3343 Belgium
19: 13729 Honda
20: 18769304 Shayne_King
21: 4913064 New_Zealand
22: 243320 KTM
23: 14948713 Jacky_Martens
24: 3343 Belgium
25: 13504069 Husqvarna_Motorcycles
26: 11867 Germany
27: 13729 Honda
28: 76972 Denmark
29: 243320 KTM
30: 14948713 Jacky_Martens
31: 18769304 Shayne_King
32: 1226329 Stefan_Johansson
33: 18769304 Shayne_King
34: 1226329 Stefan_Johansson
35: 4913064 New_Zealand
36: 13729 Honda
1: 5058739 Sweden
[pem_pos=0, final_scores=0.40387273,
final_local_scores=0.40612495, pem=-0.63135606,
similarity=19.642704]
2: 1019545 Landskrona
[pem_pos=0, final_scores=0.41429323,
final_local_scores=0.47415543, pem=0.0, similarity=18.239456]

```
3: 5058739 Sweden
    [pem_pos=0, final_scores=0.7429213,
final_local_scores=0.776484, pem=-0.2135508,
similarity=28.422867]
4: 5058739 Sweden
    [pem_pos=0, final_scores=0.71492183,
final_local_scores=0.74589884, pem=-0.63135606,
similarity=29.782234]
5: 6369246 Joël_Smets
    [pem_pos=0, final_scores=0.45604187,
final_local_scores=0.52318156, pem=0.0, similarity=19.702494]
6: 3343 Belgium
    [pem_pos=0, final_scores=0.71114784,
final_local_scores=0.7248364, pem=-0.24180195,
similarity=27.035234]
7: 3347463 Husaberg
    [pem_pos=0, final_scores=0.39765233,
final_local_scores=0.4608435, pem=0.0, similarity=17.842203]
8: 18781780 Peter_Johansson_(motorcyclist)
    [pem_pos=1, final_scores=0.3649431,
final_local_scores=0.47425687, pem=-0.4937057,
similarity=20.927332]
9: 5058739 Sweden
    [pem_pos=0, final_scores=0.77362734,
final_local_scores=0.8100257, pem=-0.2135508,
similarity=29.423817]
10: 13504069 Husqvarna_Motorcycles
    [pem_pos=0, final_scores=0.54139966,
final_local_scores=0.5727565, pem=0.0, similarity=21.18191]
11: 21148 Netherlands
    [pem_pos=0, final_scores=0.63005483,
final_local_scores=0.64779025, pem=-0.24940795,
similarity=24.777386]
12: 13729 Honda
    [pem_pos=0, final_scores=0.2749951,
final_local_scores=0.26274374, pem=-0.17204832,
similarity=12.866136]
13: 14948713 Jacky_Martens
    [pem_pos=0, final_scores=0.54728585,
final_local_scores=0.6119503, pem=0.0, similarity=22.35153]
14: 3343 Belgium
    [pem_pos=0, final_scores=0.72659636,
final_local_scores=0.74171156, pem=-0.24180195,
similarity=27.538822]
15: 13504069 Husqvarna_Motorcycles
    [pem_pos=0, final_scores=0.5712751,
final_local_scores=0.6053909, pem=0.0, similarity=22.155785]
16: 3343 Belgium
    [pem_pos=0, final_scores=0.72473717,
final_local_scores=0.73968065, pem=-0.24180195,
similarity=27.478216]
17: 243320 KTM
```

```
[pem_pos=0, final_scores=0.47594064,
final_local_scores=0.4838786, pem=0.0, similarity=18.529615]
18: 3343 Belgium
    [pem_pos=0, final_scores=0.7097097,
final_local_scores=0.72326547, pem=-0.24180195,
similarity=26.988354]
19: 13729 Honda
    [pem_pos=0, final_scores=0.35393173,
final_local_scores=0.34897006, pem=-0.17204832,
similarity=15.4393015]
20: 18769304 Shayne_King
    [pem_pos=0, final_scores=0.49158728,
final_local_scores=0.58697534, pem=0.0, similarity=21.606228]
21: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.5416755,
final_local_scores=0.6083614, pem=-0.23795462,
similarity=23.538465]
22: 243320 KTM
    [pem_pos=0, final_scores=0.5156998,
final_local_scores=0.52730954, pem=0.0, similarity=19.825682]
23: 14948713 Jacky_Martens
    [pem_pos=0, final_scores=0.5959659,
final_local_scores=0.6651258, pem=0.0, similarity=23.938393]
24: 3343 Belgium
    [pem_pos=0, final_scores=0.7225584,
final_local_scores=0.7373007, pem=-0.24180195,
similarity=27.407194]
25: 13504069 Husqvarna_Motorcycles
    [pem_pos=0, final_scores=0.55105954,
final_local_scores=0.58330846, pem=0.0, similarity=21.496801]
26: 11867 Germany
    [pem_pos=0, final_scores=0.8140498,
final_local_scores=0.8605152, pem=-0.15529114,
similarity=30.613699]
27: 13729 Honda
    [pem_pos=0, final_scores=0.32560033,
final_local_scores=0.3180223, pem=-0.17204832,
similarity=14.5157585]
28: 76972 Denmark
    [pem_pos=0, final_scores=0.5782358,
final_local_scores=0.6002425, pem=-0.26762357,
similarity=23.457525]
29: 243320 KTM
    [pem_pos=0, final_scores=0.4422999,
final_local_scores=0.44713122, pem=0.0, similarity=17.433]
30: 14948713 Jacky_Martens
    [pem_pos=0, final_scores=0.50123364,
final_local_scores=0.5616452, pem=0.0, similarity=20.850327]
31: 18769304 Shayne_King
    [pem_pos=0, final_scores=0.46639472,
final_local_scores=0.5594563, pem=0.0, similarity=20.785004]
32: 1226329 Stefan_Johansson
```

```

    [pem_pos=6, final_scores=0.54007745,
final_local_scores=0.6122568, pem=-2.016291,
similarity=33.32558]
33: 18769304 Shayne_King
    [pem_pos=0, final_scores=0.53059554,
final_local_scores=0.6295859, pem=0.0, similarity=22.877813]
34: 1226329 Stefan_Johansson
    [pem_pos=6, final_scores=0.5309802,
final_local_scores=0.60231954, pem=-2.016291,
similarity=33.029034]
35: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.5289181,
final_local_scores=0.5944258, pem=-0.23795462,
similarity=23.122599]
36: 13729 Honda
    [pem_pos=0, final_scores=0.34799367,
final_local_scores=0.34248358, pem=-0.17204832,
similarity=15.245731]
```

Sample: b'315_MOTOCROSS'

MOTOCROSS - [1 1 GERMAN 1] 125CC GRAND PRIX RESULTS . [2 2 HO
LZGERLINGEN 2] , [3 3 Germany 3] 1996-08-25 Leading results
in the [4 4 German 4] 125cc motocross Grand Prix on Sunday :
First race 1. [5 5 Sebastien Tortelli 5] ([6 6 France 6]) [7
7 Kawasaki 7] 2. Bob Moore ([8 8 U.S. 8]) [9 9 Yamaha 9] 3.
Luigi Seguy ([10 10 France 10]) TM 4. Andi Kanstinger ([11
11 Germany 11]) [12 12 Honda 12] 5. Nicolas Charlier ([13 13
France 13]) [14 14 Kawasaki 14] 6. Erik Camerlengo ([15 15 I
taly 15]) [16 16 Yamaha 16] Second race 1. [17 17 Tortelli 17
] 2. Moore 3. Alex Belometti ([18 18 Italy 18]) [19 19 Honda
19] 4. Frederic Vialle ([20 20 France 20]) [21 21 Yamaha 21]
5. Collin Dugmore ([22 22 South Africa 22]) [23 23 Honda 23]
6. Camerlengo Overall on day : 1. [24 24 Tortelli 24] 40
points 2. Moore 34 3. Seguy 24 4. Vialle 22 5. Camerlengo 20
6. Belometti 19 Final world championship standings : 1. [25 25
Tortelli 25] 432 points 2. Paul Malin ([26 26 Britain 26]) [
27 27 Yamaha 27] 317 3. Vialle 293 4. Seguy 192 5. Michele
Fanton ([28 28 Italy 28]) [29 29 Kawasaki 29] 160 6.
Dugmore 152

1: 11867 Germany
2: 4430749 Holzgerlingen
3: 11867 Germany
4: 11867 Germany
5: 4632360 Sébastien_Tortelli
6: 5843419 France
7: 15398376 Kawasaki_Motors_Racing
8: 3434750 United_States
9: 332378 Yamaha_Motor_Company
10: 5843419 France
11: 11867 Germany
12: 6527836 Honda_Racing_Corporation
13: 5843419 France

```
14: 15398376 Kawasaki_Motors_Racing
15: 14532 Italy
16: 332378 Yamaha_Motor_Company
17: 4632360 Sébastien_Tortelli
18: 14532 Italy
19: 6527836 Honda_Racing_Corporation
20: 5843419 France
21: 332378 Yamaha_Motor_Company
22: 17416221 South_Africa
23: 6527836 Honda_Racing_Corporation
24: 4632360 Sébastien_Tortelli
25: 4632360 Sébastien_Tortelli
26: 31717 United_Kingdom
27: 332378 Yamaha_Motor_Company
28: 14532 Italy
29: 15398376 Kawasaki_Motors_Racing
1: 11867 Germany
    [pem_pos=1, final_scores=0.5550615,
final_local_scores=0.5814329, pem=-1.0465823,
similarity=27.132307]
2: 4430749 Holzgerlingen
    [pem_pos=0, final_scores=0.27120692,
final_local_scores=0.400679, pem=0.0, similarity=16.046774]
3: 11867 Germany
    [pem_pos=0, final_scores=0.86934537,
final_local_scores=0.92474014, pem=-0.15529114,
similarity=32.5303]
4: 11867 Germany
    [pem_pos=1, final_scores=0.82204753,
final_local_scores=0.8730745, pem=-1.0465823,
similarity=35.835476]
5: 4632360 Sébastien_Tortelli
    [pem_pos=0, final_scores=0.6703869,
final_local_scores=0.712527, pem=0.0, similarity=25.35294]
6: 5843419 France
    [pem_pos=0, final_scores=0.73045594,
final_local_scores=0.7636767, pem=-0.16240834,
similarity=27.76255]
7: 15398376 Kawasaki_Motors_Racing
    [pem_pos=7, final_scores=0.39339048,
final_local_scores=0.4385957, pem=-3.3328056,
similarity=35.302597]
8: 3434750 United_States
    [pem_pos=0, final_scores=0.4838677,
final_local_scores=0.5146217, pem=-0.14075686,
similarity=20.21251]
9: 332378 Yamaha_Motor_Company
    [pem_pos=1, final_scores=0.53843343,
final_local_scores=0.5574667, pem=-0.7066168,
similarity=24.568323]
10: 5843419 France
    [pem_pos=0, final_scores=0.75529104,
```

```
final_local_scores=0.7908052, pem=-0.16240834,
similarity=28.572119]
11: 11867 Germany
    [pem_pos=0, final_scores=0.8425539,
final_local_scores=0.8954746, pem=-0.15529114,
similarity=31.656958]
12: 6527836 Honda_Racing_Corporation
    [pem_pos=5, final_scores=0.27444565,
final_local_scores=0.30685842, pem=-3.9711916,
similarity=34.842937]
13: 5843419 France
    [pem_pos=0, final_scores=0.71809715,
final_local_scores=0.75017655, pem=-0.16240834,
similarity=27.359678]
14: 15398376 Kawasaki_Motors_Racing
    [pem_pos=7, final_scores=0.4336096,
final_local_scores=0.48252898, pem=-3.3328056,
similarity=36.613655]
15: 14532 Italy
    [pem_pos=0, final_scores=0.7505606,
final_local_scores=0.7936347, pem=-0.20612453,
similarity=28.894293]
16: 332378 Yamaha_Motor_Company
    [pem_pos=1, final_scores=0.5020556,
final_local_scores=0.51772946, pem=-0.7066168,
similarity=23.382486]
17: 4632360 Sébastien_Tortelli
    [pem_pos=0, final_scores=0.8822263,
final_local_scores=0.9520249, pem=0.0, similarity=32.500034]
18: 14532 Italy
    [pem_pos=0, final_scores=0.67811483,
final_local_scores=0.7144987, pem=-0.20612453,
similarity=26.532717]
19: 6527836 Honda_Racing_Corporation
    [pem_pos=5, final_scores=0.40318692,
final_local_scores=0.44402397, pem=-3.9711916,
similarity=38.93623]
20: 5843419 France
    [pem_pos=0, final_scores=0.57385534,
final_local_scores=0.5926143, pem=-0.16240834,
similarity=22.657707]
21: 332378 Yamaha_Motor_Company
    [pem_pos=1, final_scores=0.5364959,
final_local_scores=0.55535024, pem=-0.7066168,
similarity=24.505165]
22: 17416221 South_Africa
    [pem_pos=0, final_scores=0.5544379,
final_local_scores=0.63241225, pem=-0.26251382,
similarity=24.389748]
23: 6527836 Honda_Racing_Corporation
    [pem_pos=5, final_scores=0.3887732,
final_local_scores=0.42827916, pem=-3.9711916,
```

```
similarity=38.466373]
24: 4632360 Sébastien_Tortelli
    [pem_pos=0, final_scores=0.867762,
final_local_scores=0.9362249, pem=0.0, similarity=32.02853]
25: 4632360 Sébastien_Tortelli
    [pem_pos=0, final_scores=0.82847875,
final_local_scores=0.8933138, pem=0.0, similarity=30.74798]
26: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5983271,
final_local_scores=0.6432187, pem=-0.61114657,
similarity=26.60815]
27: 332378 Yamaha_Motor_Company
    [pem_pos=1, final_scores=0.4803095,
final_local_scores=0.4939751, pem=-0.7066168,
similarity=22.673607]
28: 14532 Italy
    [pem_pos=0, final_scores=0.7361495,
final_local_scores=0.77789277, pem=-0.20612453,
similarity=28.424522]
29: 15398376 Kawasaki_Motors_Racing
    [pem_pos=7, final_scores=0.29385996,
final_local_scores=0.3313952, pem=-3.3328056,
similarity=32.103523]
```

Sample: b'316_MOTOR'

MOTOR RACING - LEADING PLACINGS IN POKKA 1,000 KM RACE . [1 1 SUZUKA 1] , [2 2 Japan 2] 1996-08-25 Leading placings in Sunday 's Pokka 1,000 km motor race , seventh round of the International Endurance GT championship : 1. Ray Belim ([3 3 Britain 3]) / [4 4 James Weaver 4] ([5 5 Britain 5]) / J.J.Lehto ([6 6 Finland 6]) Gulf McLaren FI GTR 171 laps - 6 hours 18 minutes 48.637 seconds (average speed 158.82 kph) 2. Anders Olofsson ([7 7 Sweden 7]) / Luciano della Noce ([8 8 Italy 8]) Ennea Ferrari F40 170 laps 3. Andy Ballace ([9 9 Britain 9]) / [10 10 Olivier Grouillard 10] ([11 11 France 11]) Harrods McLaren FI GTR 169 4. [12 12 Thomas Bscher 12] ([13 13 Germany 13]) / [14 14 Peter Kox 14] ([15 15 Netherlands 15]) West McLaren F1 GTR 168 5. Fabien Giroix ([16 16 France 16]) / [17 17 Jean-Denis Deletraz 17] ([18 18 Switzerland 18]) Muller McLaren F1 GTR 167 6. Lindsay Owen-Jones ([19 19 Britain 19]) / [20 20 Pierre-Henri Raphanel 20] ([21 21 France 21]) / [22 22 David Brabham 22] ([23 23 Australia 23]) Gulf McLaren F1 GTR 167 7. [24 24 Jean-Marc Gounon 24] ([25 25 France 25]) / [26 26 Eric Bernard 26] ([27 27 France 27]) / [28 28 Paul Belmondo 28] ([29 29 France 29]) Ennea Ferrari F40 167 8. Bruno Eichmann ([30 30 Germany 30]) / Gerd Ruch ([31 31 Germany 31]) / [32 32 Ralf Kelleners 32] ([33 33 Germany 33]) GT2 Roock Porsche 911 164 9. [34 34 Stephane Ortelli 34] ([35 35 France 35]) / [36 36 Bob Wollek 36] ([37 37 France 37]) / [38 38 Franz Konrad 38] ([39 39 Austria 39]) GT2 Konrad Porsche 911 164 10. [40 40 Cor Euser 40] ([41 41 Netherlands 41]) / H.

Wada ([42 42 Japan 42]) / N. Furuya ([43 43 Japan 43])
GT2 Marcos LM600 162 Fastest lap : [44 44 Gounon 44] , 2
minutes 03.684 seconds (170.680 kph) Championship standings
after seven rounds : 1. Belim , Weaver 156 points 2.
Eichmann , Ruch 116 3. [45 45 Bscher 45] 112 4. [46 46 Gounon
46] , 1 Bernard , [47 47 Belmondo 47] 98 5. Olofsson , della
Noce 93 6. Owen-Jones , [48 48 Raphanel 48] 82
1: 748984 Suzuka_Circuit
2: 15573 Japan
3: 31717 United_Kingdom
4: 5691770 James_Weaver_(racing_driver)
5: 31717 United_Kingdom
6: 10577 Finland
7: 5058739 Sweden
8: 14532 Italy
9: 31717 United_Kingdom
10: 1226512 Olivier_Grouillard
11: 5843419 France
12: 1877657 Thomas_Bscher
13: 11867 Germany
14: 12576820 Peter_Kox
15: 21148 Netherlands
16: 5843419 France
17: 1233974 Jean-Denis_Délétraz
18: 26748 Switzerland
19: 31717 United_Kingdom
20: 1237455 Pierre-Henri_Raphanel
21: 5843419 France
22: 1148454 David_Brabham
23: 4689264 Australia
24: 1234917 Jean-Marc_Gounon
25: 5843419 France
26: 1219780 Éric_Bernard
27: 5843419 France
28: 1235361 Paul_Belmondo
29: 5843419 France
30: 11867 Germany
31: 11867 Germany
32: 5166757 Ralf_Kelleners
33: 11867 Germany
34: 5075311 Stéphane_Ortelli
35: 5843419 France
36: 3650116 Bob_Wollek
37: 5843419 France
38: 5166840 Franz_Konrad_(racing_driver)
39: 26964606 Austria
40: 23507562 Cor_Euser
41: 21148 Netherlands
42: 15573 Japan
43: 15573 Japan
44: 1234917 Jean-Marc_Gounon
45: 1877657 Thomas_Bscher

46: 1234917 Jean-Marc_Gounon
47: 1235361 Paul_Belmondo
48: 1237455 Pierre-Henri_Raphanel
1: 1228226 Bernard_Collomb
[winner: 1219780 Éric_Bernard, pem_pos=0,
final_scores=-0.17767991, final_local_scores=-0.16966361,
pem=0.0, similarity=-0.9733877]
RECALL MISS: 1228226 Bernard_Collomb]
1: 748984 Suzuka_Circuit
[pem_pos=0, final_scores=0.19957675,
final_local_scores=0.23937243, pem=-0.51707464,
similarity=14.044996]
2: 15573 Japan
[pem_pos=0, final_scores=0.7378127,
final_local_scores=0.7869863, pem=-0.13549899,
similarity=28.311817]
3: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5767094,
final_local_scores=0.59015805, pem=-0.61114657,
similarity=25.024715]
4: 5691770 James_Weaver_(racing_driver)
[pem_pos=0, final_scores=0.3500268,
final_local_scores=0.4456643, pem=0.0, similarity=17.389225]
5: 31717 United_Kingdom
[pem_pos=0, final_scores=0.567034,
final_local_scores=0.5795892, pem=-0.61114657,
similarity=24.70932]
6: 10577 Finland
[pem_pos=0, final_scores=0.74054444,
final_local_scores=0.7745227, pem=-0.2445894,
similarity=28.53313]
7: 5058739 Sweden
[pem_pos=0, final_scores=0.7673295,
final_local_scores=0.8058811, pem=-0.2135508,
similarity=29.300135]
8: 14532 Italy
[pem_pos=0, final_scores=0.72457933,
final_local_scores=0.74998593, pem=-0.20612453,
similarity=27.591726]
9: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5931578,
final_local_scores=0.6081255, pem=-0.61114657,
similarity=25.5609]
10: 1226512 Olivier_Grouillard
[pem_pos=0, final_scores=0.6769509,
final_local_scores=0.76772875, pem=0.0, similarity=27.00027]
11: 5843419 France
[pem_pos=0, final_scores=0.8070261,
final_local_scores=0.8083134, pem=-0.16240834,
similarity=29.094599]
12: 1877657 Thomas_Bscher
[pem_pos=0, final_scores=0.37603152,

```
final_local_scores=0.4929347, pem=0.0, similarity=18.799868]
13: 11867 Germany
    [pem_pos=0, final_scores=0.8702658,
final_local_scores=0.8934489, pem=-0.15529114,
similarity=31.596508]
14: 12576820 Peter_Kox
    [pem_pos=0, final_scores=0.49498147,
final_local_scores=0.5809653, pem=0.0, similarity=21.426876]
15: 21148 Netherlands
    [pem_pos=0, final_scores=0.62637347,
final_local_scores=0.63171875, pem=-0.24940795,
similarity=24.297781]
16: 5843419 France
    [pem_pos=0, final_scores=0.7962244,
final_local_scores=0.7965143, pem=-0.16240834,
similarity=28.742489]
17: 1233974 Jean-Denis_Délétraz
    [pem_pos=0, final_scores=0.62614053,
final_local_scores=0.7197052, pem=0.0, similarity=25.567152]
18: 26748 Switzerland
    [pem_pos=0, final_scores=0.6996723,
final_local_scores=0.70505464, pem=-0.18598008,
similarity=26.141338]
19: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5951226,
final_local_scores=0.61027163, pem=-0.61114657,
similarity=25.624947]
20: 1237455 Pierre-Henri_Raphanel
    [pem_pos=0, final_scores=0.57217973,
final_local_scores=0.66775095, pem=0.0, similarity=24.016733]
21: 5843419 France
    [pem_pos=0, final_scores=0.7437975,
final_local_scores=0.7392458, pem=-0.16240834,
similarity=27.033482]
22: 1148454 David_Brabham
    [pem_pos=0, final_scores=0.5944022,
final_local_scores=0.6770876, pem=0.0, similarity=24.295357]
23: 4689264 Australia
    [pem_pos=0, final_scores=0.6803642,
final_local_scores=0.72582924, pem=-0.17168477,
similarity=26.683554]
24: 1234917 Jean-Marc_Gounon
    [pem_pos=0, final_scores=0.54596317,
final_local_scores=0.64549255, pem=0.0, similarity=23.352499]
25: 5843419 France
    [pem_pos=0, final_scores=0.7579246,
final_local_scores=0.7546776, pem=-0.16240834,
similarity=27.494]
26: 1219780 Éric_Bernard
    [pem_pos=0, final_scores=0.5756078,
final_local_scores=0.66032666, pem=0.0, similarity=23.795177]
27: 5843419 France
```

```
[pem_pos=0, final_scores=0.72209996,
final_local_scores=0.7155445, pem=-0.16240834,
similarity=26.326189]
28: 1235361 Paul_Belmondo
    [pem_pos=0, final_scores=0.47594774,
final_local_scores=0.5408574, pem=0.0, similarity=20.229977]
29: 5843419 France
    [pem_pos=0, final_scores=0.7436163,
final_local_scores=0.7390479, pem=-0.16240834,
similarity=27.027576]
30: 11867 Germany
    [pem_pos=0, final_scores=0.8120218,
final_local_scores=0.8298261, pem=-0.15529114,
similarity=29.697878]
31: 11867 Germany
    [pem_pos=0, final_scores=0.7301654,
final_local_scores=0.74041045, pem=-0.15529114,
similarity=27.029535]
32: 5166757 Ralf_Kelleners
    [pem_pos=0, final_scores=0.53486866,
final_local_scores=0.6457329, pem=0.0, similarity=23.35967]
33: 11867 Germany
    [pem_pos=0, final_scores=0.80574745,
final_local_scores=0.82297236, pem=-0.15529114,
similarity=29.493347]
34: 5075311 Stéphane_Ortelli
    [pem_pos=0, final_scores=0.46300656,
final_local_scores=0.5901015, pem=0.0, similarity=21.699518]
35: 5843419 France
    [pem_pos=0, final_scores=0.7759777,
final_local_scores=0.77439785, pem=-0.16240834,
similarity=28.08249]
36: 3650116 Bob_Wollek
    [pem_pos=0, final_scores=0.5907413,
final_local_scores=0.68648523, pem=0.0, similarity=24.575802]
37: 5843419 France
    [pem_pos=0, final_scores=0.7647019,
final_local_scores=0.76208067, pem=-0.16240834,
similarity=27.714922]
38: 5166840 Franz_Konrad_(racing_driver)
    [pem_pos=1, final_scores=0.4795189,
final_local_scores=0.5727852, pem=-1.5404451,
similarity=29.559946]
39: 26964606 Austria
    [pem_pos=0, final_scores=0.6369765,
final_local_scores=0.64127797, pem=-0.25262418,
similarity=24.600536]
40: 23507562 Cor_Euser
    [pem_pos=0, final_scores=0.52241117,
final_local_scores=0.62791264, pem=0.0, similarity=22.827879]
41: 21148 Netherlands
    [pem_pos=0, final_scores=0.6489026,
```

```
final_local_scores=0.65632844, pem=-0.24940795,
similarity=25.032183]
42: 15573 Japan
    [pem_pos=0, final_scores=0.6520491,
final_local_scores=0.69330263, pem=-0.13549899,
similarity=25.51611]
43: 15573 Japan
    [pem_pos=0, final_scores=0.7143021,
final_local_scores=0.76130456, pem=-0.13549899,
similarity=27.545424]
44: 1234917 Jean-Marc_Gounon
    [pem_pos=0, final_scores=0.45132262,
final_local_scores=0.5421122, pem=0.0, similarity=20.267422]
45: 1877657 Thomas_Bscher
    [pem_pos=0, final_scores=0.50776494,
final_local_scores=0.6368334, pem=0.0, similarity=23.094093]
46: 1234917 Jean-Marc_Gounon
    [pem_pos=0, final_scores=0.4517784,
final_local_scores=0.54261005, pem=0.0, similarity=20.28228]
47: 1235361 Paul_Belmondo
    [pem_pos=0, final_scores=0.5654177,
final_local_scores=0.64573675, pem=0.0, similarity=23.359785]
48: 1237455 Pierre-Henri_Raphanel
    [pem_pos=0, final_scores=0.5595689,
final_local_scores=0.6539755, pem=0.0, similarity=23.605646]
```

Sample: b'317 ATHLETICS'

ATHLETICS - LEADING RESULTS AT [1 1 SHEFFIELD 1]
INTERNATIONAL MEETING . [2 2 SHEFFIELD 2] , [3 3 England 3]
1996-08-25 Leading results at an international meeting on
Sunday : Women 's triple jump 1. [4 4 Sarka Kasparkova 4] ([5
5 Czech Republic 5]) 14.84 metres 2. [6 6 Ashia Hansen 6] ([
7 7 Britain 7]) 14.78 3. Rodica Matescu ([8 8 Romania 8])
14.18 Women 's 400 metres hurdles 1. [9 9 Deon Hemmings 9] ([
10 10 Jamaica 10]) 55.13 seconds 2. Anne Marken ([11 11 Bel
gium 11]) 55.90 3. Susan Smith ([12 12 Ireland 12]) 56.00
Women 's javelin 1. [13 13 Isel Lopez 13] ([14 14 Cuba 14])
61.36 2. [15 15 Louise McPaul 15] ([16 16 Australia 16])
60.66 3. [17 17 Silke Renk 17] ([18 18 Germany 18]) 60.66
Women 's 200 metres 1. [19 19 Cathy Freeman 19] ([20 20 Aust
ralia 20]) 22.53 2. [21 21 Falilat Ogunkoya 21] ([22 22 Nig
eria 22]) 22.58 3. [23 23 Juliet Cuthbert 23] ([24 24 Jamai
ca 24]) 22.77 100 metres hurdles 1. Dionne Rose ([25 25 Jam
aica 25]) 12.83 2. [26 26 Michelle Freeman 26] ([27 27 Jama
ica 27]) 12.91 3. [28 28 Gillian Russell 28] ([29 29 Jamaica
29]) 12.95 Women 's 800 metres 1. [30 30 Charmaine Crooks 30]
([31 31 Canada 31]) two minutes 00.42 seconds 2. [32 32 Inez
Turner 32] ([33 33 Jamaica 33]) 2:01.98 3. [34 34 Margaret C
rowley 34] ([35 35 Australia 35]) 2:02.40 Men 's pole vault
1. Trond Bathel ([36 36 Norway 36]) 5.60 2. [37 37 Pat Mans
on 37] ([38 38 U.S. 38]) 5.60 3. [39 39 Tim Lobinger 39] ([
40 40 Germany 40]) 5.50 Men 's javelin 1. [41 41 Tom Pukstys

41] ([42 42 U.S. 42]) 86.82 2. [43 43 Steve Backley 43] ([4
4 44 Britain 44]) 82.20 3. [45 45 Nick Nieland 45] ([46 46 B
ritain 46]) 81.12 Women 's 400 metres 1. Marcel Malone ([47
47 U.S. 47]) 51.50 2. [48 48 Kim Graham 48] ([49 49 U.S. 49]
) 52.17 3. [50 50 Phylis Smith 50] ([51 51 Britain 51])
52.53 Men 's 200 metres 1. [52 52 Jeff Williams 52] ([53 53 U
.S. 53]) 20.45 2. Doug Turner ([54 54 Britain 54]) 20.48
3. [55 55 John Regis 55] ([56 56 Britain 56]) 20.63 Men 's
high jump 1. [57 57 Charles Austin 57] ([58 58 U.S. 58])
2.30 2. [59 59 Tim Forsyth 59] ([60 60 Australia 60]) 2.30
3. [61 61 Patrik Sjoberg 61] ([62 62 Sweden 62]) 2.25 Men
's 800 metres 1. Verbjorn Rodal ([63 63 Norway 63]) 1:44.93
2. [64 64 Benson Koech 64] ([65 65 Kenya 65]) 1:45.96 3.
Vincent Malakwen ([66 66 Kenya 66]) 1:46.18 Men 's mile 1. [
67 67 William Tanui 67] ([68 68 Kenya 68]) 3:54.57 2. [69 69
John Mayock 69] ([70 70 Britain 70]) 3:54.60 3. Tony
Whiteman ([71 71 Britain 71]) 3:54.87 Men 's 400 metres 1. [
72 72 Roger Black 72] ([73 73 Britain 73]) 45.05 2. [74 74 M
ark Richardson 74] ([75 75 Britain 75]) 45.38 3. [76 76 Der
ek Mills 76] ([77 77 U.S. 77]) 45.48 Men 's 100 metres 1. [
78 78 Osmond Ezinwa 78] ([79 79 Nigeria 79]) 10.06 2. Ian
Mackie ([80 80 Britain 80]) 10.17 3. [81 81 Linford Christie
81] ([82 82 Britain 82]) 10.19

1: 88885 Sheffield
2: 88885 Sheffield
3: 9316 England
4: 2390745 Šárka_Kašpárková
5: 5321 Czech_Republic
6: 1709004 Ashia_Hansen
7: 31717 United_Kingdom
8: 25445 Romania
9: 6397642 Deon_Hemmings
10: 15660 Jamaica
11: 3343 Belgium
12: 14560 Republic_of_Ireland
13: 9831365 Isel_López
14: 5042481 Cuba
15: 3874782 Louise_Currey
16: 4689264 Australia
17: 3910212 Silke_Renk
18: 11867 Germany
19: 66002 Cathy_Freeman
20: 4689264 Australia
21: 2489403 Falilat_Ogunkoya
22: 21383 Nigeria
23: 6756205 Juliet_Cuthbert
24: 15660 Jamaica
25: 15660 Jamaica
26: 6829072 Michelle_Freeman
27: 15660 Jamaica
28: 6855453 Gillian_Russell
29: 15660 Jamaica

30: 6734797 Charmaine_Crooks
31: 5042916 Canada
32: 2736349 Inez_Turner
33: 15660 Jamaica
34: 24234976 Margaret_Crowley_(athlete)
35: 4689264 Australia
36: 21241 Norway
37: 6245295 Pat_Manson
38: 3434750 United_States
39: 3917330 Tim_Lobinger
40: 11867 Germany
41: 14086432 Tom_Pukstys
42: 3434750 United_States
43: 412057 Steve_Backley
44: 31717 United_Kingdom
45: 4856056 Nick_Nieland
46: 31717 United_Kingdom
47: 3434750 United_States
48: 2645833 Kim_Graham
49: 3434750 United_States
50: 2679515 Phylis_Smith
51: 31717 United_Kingdom
52: 27446348 Jeff_Williams_(athlete)
53: 3434750 United_States
54: 31717 United_Kingdom
55: 3373156 John_Regis_(athlete)
56: 31717 United_Kingdom
57: 3909733 Charles_Austin
58: 3434750 United_States
59: 3450852 Tim_Forsyth
60: 4689264 Australia
61: 2254671 Patrik_Sjöberg
62: 5058739 Sweden
63: 21241 Norway
64: 3552049 Benson_Koech
65: 188171 Kenya
66: 188171 Kenya
67: 815988 William_Tanui
68: 188171 Kenya
69: 13182495 John_Mayock
70: 31717 United_Kingdom
71: 31717 United_Kingdom
72: 805864 Roger_Black
73: 31717 United_Kingdom
74: 7652893 Mark_Richardson_(athlete)
75: 31717 United_Kingdom
76: 1676095 Derek_Mills
77: 3434750 United_States
78: 2736304 Osmond_Ezinwa
79: 21383 Nigeria
80: 31717 United_Kingdom
81: 67118 Linford_Christie

```
82: 31717 United_Kingdom
1: 88885 Sheffield
    [pem_pos=0, final_scores=0.3681668,
final_local_scores=0.36346114, pem=-0.25810194,
similarity=16.339716]
2: 88885 Sheffield
    [pem_pos=0, final_scores=0.6929846,
final_local_scores=0.71827525, pem=-0.25810194,
similarity=26.928078]
3: 9316 England
    [pem_pos=0, final_scores=0.6461092,
final_local_scores=0.64641494, pem=-0.19053856,
similarity=24.416203]
4: 2390745 Šárka_Kašpárková
    [pem_pos=0, final_scores=0.85235935,
final_local_scores=1.0062199, pem=0.0, similarity=34.11732]
5: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.6239944,
final_local_scores=0.65212077, pem=-0.27794555,
similarity=25.06181]
6: 1709004 Ashia_Hansen
    [pem_pos=0, final_scores=0.73149496,
final_local_scores=0.84839094, pem=0.0, similarity=29.40739]
7: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.6328251,
final_local_scores=0.6081622, pem=-0.61114657,
similarity=25.561996]
8: 25445 Romania
    [pem_pos=0, final_scores=0.72375745,
final_local_scores=0.7645327, pem=-0.21241479,
similarity=28.060038]
9: 6397642 Deon_Hemmings
    [pem_pos=0, final_scores=0.6717698,
final_local_scores=0.8294831, pem=0.0, similarity=28.843143]
10: 15660 Jamaica
    [pem_pos=0, final_scores=0.4938106,
final_local_scores=0.46739036, pem=-0.33295843,
similarity=19.848253]
11: 3343 Belgium
    [pem_pos=0, final_scores=0.7259947,
final_local_scores=0.74013436, pem=-0.24180195,
similarity=27.491756]
12: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.5752862,
final_local_scores=0.60552996, pem=-1.5040284,
similarity=30.339075]
13: 9831365 Isel_López
    [pem_pos=0, final_scores=0.5248253,
final_local_scores=0.7375603, pem=0.0, similarity=26.099983]
14: 5042481 Cuba
    [pem_pos=0, final_scores=0.52113295,
final_local_scores=0.5275344, pem=-0.23907764,
```

```
similarity=21.132534]
15: 3874782 Louise_Currey
    [pem_pos=0, final_scores=0.58295256,
final_local_scores=0.7135434, pem=0.0, similarity=25.383272]
16: 4689264 Australia
    [pem_pos=0, final_scores=0.72775465,
final_local_scores=0.7311544, pem=-0.17168477,
similarity=26.842466]
17: 3910212 Silke_Renk
    [pem_pos=0, final_scores=0.7109447,
final_local_scores=0.8733855, pem=0.0, similarity=30.153276]
18: 11867 Germany
    [pem_pos=0, final_scores=0.84943277,
final_local_scores=0.88170344, pem=-0.15529114,
similarity=31.246]
19: 66002 Cathy_Freeman
    [pem_pos=0, final_scores=0.6962318,
final_local_scores=0.7775871, pem=0.0, similarity=27.294464]
20: 4689264 Australia
    [pem_pos=0, final_scores=0.73316246,
final_local_scores=0.7370616, pem=-0.17168477,
similarity=27.018751]
21: 2489403 Falilat_Ogunkoya
    [pem_pos=0, final_scores=0.7872669,
final_local_scores=0.9351439, pem=0.0, similarity=31.996273]
22: 21383 Nigeria
    [pem_pos=0, final_scores=0.6407305,
final_local_scores=0.6719227, pem=-0.31455088,
similarity=25.851803]
23: 6756205 Juliet_Cuthbert
    [pem_pos=0, final_scores=0.78247195,
final_local_scores=0.92874473, pem=0.0, similarity=31.805307]
24: 15660 Jamaica
    [pem_pos=0, final_scores=0.49661398,
final_local_scores=0.4704526, pem=-0.33295843,
similarity=19.939636]
25: 15660 Jamaica
    [pem_pos=0, final_scores=0.51453286,
final_local_scores=0.49002624, pem=-0.33295843,
similarity=20.523752]
26: 6829072 Michelle_Freeman
    [pem_pos=0, final_scores=0.6328899,
final_local_scores=0.7886936, pem=0.0, similarity=27.625904]
27: 15660 Jamaica
    [pem_pos=0, final_scores=0.51994133,
final_local_scores=0.4959342, pem=-0.33295843,
similarity=20.700058]
28: 6855453 Gillian_Russell
    [pem_pos=0, final_scores=0.5903002,
final_local_scores=0.7154228, pem=0.0, similarity=25.439356]
29: 15660 Jamaica
    [pem_pos=0, final_scores=0.5280333,
```

```
final_local_scores=0.5047735, pem=-0.33295843,
similarity=20.96384]
30: 6734797 Charmaine_Crooks
    [pem_pos=0, final_scores=0.44877356,
final_local_scores=0.5484166, pem=0.0, similarity=20.455559]
31: 5042916 Canada
    [pem_pos=0, final_scores=0.73366106,
final_local_scores=0.73501235, pem=-0.11662793,
similarity=26.658188]
32: 2736349 Inez_Turner
    [pem_pos=0, final_scores=0.587533,
final_local_scores=0.7318549, pem=0.0, similarity=25.929724]
33: 15660 Jamaica
    [pem_pos=0, final_scores=0.5005846,
final_local_scores=0.47478992, pem=-0.33295843,
similarity=20.06907]
34: 24234976 Margaret_Crowley_(athlete)
    [pem_pos=0, final_scores=0.5199216,
final_local_scores=0.65552336, pem=0.0, similarity=23.651838]
35: 4689264 Australia
    [pem_pos=0, final_scores=0.7122614,
final_local_scores=0.7142303, pem=-0.17168477,
similarity=26.33742]
36: 21241 Norway
    [pem_pos=0, final_scores=0.7726621,
final_local_scores=0.82292044, pem=-0.22668563,
similarity=29.88005]
37: 6245295 Pat_Manson
    [pem_pos=0, final_scores=0.6100961,
final_local_scores=0.7527021, pem=0.0, similarity=26.551846]
38: 3434750 United_States
    [pem_pos=0, final_scores=0.49884653,
final_local_scores=0.4706518, pem=-0.14075686,
similarity=18.900358]
39: 3917330 Tim_Lobinger
    [pem_pos=0, final_scores=0.6360818,
final_local_scores=0.77051485, pem=0.0, similarity=27.083412]
40: 11867 Germany
    [pem_pos=0, final_scores=0.8541656,
final_local_scores=0.8868734, pem=-0.15529114,
similarity=31.400284]
41: 14086432 Tom_Pukstys
    [pem_pos=0, final_scores=0.47132862,
final_local_scores=0.59688205, pem=0.0, similarity=21.901863]
42: 3434750 United_States
    [pem_pos=0, final_scores=0.5168787,
final_local_scores=0.49034923, pem=-0.14075686,
similarity=19.48817]
43: 412057 Steve_Backley
    [pem_pos=0, final_scores=0.57906586,
final_local_scores=0.66809577, pem=0.0, similarity=24.027023]
44: 31717 United_Kingdom
```

```
[pem_pos=0, final_scores=0.61645836,
final_local_scores=0.59028405, pem=-0.61114657,
similarity=25.028477]
45: 4856056 Nick_Nieland
    [pem_pos=0, final_scores=0.5428683,
final_local_scores=0.6761314, pem=0.0, similarity=24.266823]
46: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.6190535,
final_local_scores=0.59311885, pem=-0.61114657,
similarity=25.113071]
47: 3434750 United_States
    [pem_pos=0, final_scores=0.49559897,
final_local_scores=0.46710438, pem=-0.14075686,
similarity=18.794498]
48: 2645833 Kim_Graham
    [pem_pos=0, final_scores=0.57570755,
final_local_scores=0.7098334, pem=0.0, similarity=25.272556]
49: 3434750 United_States
    [pem_pos=0, final_scores=0.50275284,
final_local_scores=0.4749189, pem=-0.14075686,
similarity=19.027697]
50: 2679515 Phylis_Smith
    [pem_pos=0, final_scores=0.7709215,
final_local_scores=0.90887195, pem=0.0, similarity=31.212265]
51: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.57086957,
final_local_scores=0.5404852, pem=-0.61114657,
similarity=23.54238]
52: 27446348 Jeff_Williams_(athlete)
    [pem_pos=5, final_scores=0.32753164,
final_local_scores=0.41607982, pem=-1.5652602,
similarity=25.018494]
53: 3434750 United_States
    [pem_pos=0, final_scores=0.46339446,
final_local_scores=0.43192583, pem=-0.14075686,
similarity=17.744698]
54: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.62482315,
final_local_scores=0.5994213, pem=-0.61114657,
similarity=25.30115]
55: 3373156 John_Regis_(athlete)
    [pem_pos=0, final_scores=0.65859014,
final_local_scores=0.75832194, pem=0.0, similarity=26.719553]
56: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.58161736,
final_local_scores=0.55222553, pem=-0.61114657,
similarity=23.892733]
57: 3909733 Charles_Austin
    [pem_pos=0, final_scores=0.53879064,
final_local_scores=0.62481606, pem=0.0, similarity=22.73547]
58: 3434750 United_States
    [pem_pos=0, final_scores=0.5145275,
```

```
final_local_scores=0.48778087, pem=-0.14075686,
similarity=19.411524]
59: 3450852 Tim_Forsyth
    [pem_pos=0, final_scores=0.61024487,
final_local_scores=0.7161496, pem=0.0, similarity=25.461046]
60: 4689264 Australia
    [pem_pos=0, final_scores=0.7337474,
final_local_scores=0.73770064, pem=-0.17168477,
similarity=27.03782]
61: 2254671 Patrik_Sjöberg
    [pem_pos=0, final_scores=0.6588997,
final_local_scores=0.8191513, pem=0.0, similarity=28.53482]
62: 5058739 Sweden
    [pem_pos=0, final_scores=0.7523654,
final_local_scores=0.7786472, pem=-0.2135508,
similarity=28.48742]
63: 21241 Norway
    [pem_pos=0, final_scores=0.7825504,
final_local_scores=0.833722, pem=-0.22668563,
similarity=30.202393]
64: 3552049 Benson_Koech
    [pem_pos=0, final_scores=0.48624313,
final_local_scores=0.63549864, pem=0.0, similarity=23.05426]
65: 188171 Kenya
    [pem_pos=0, final_scores=0.61847836,
final_local_scores=0.65164316, pem=-0.27389407,
similarity=25.025522]
66: 188171 Kenya
    [pem_pos=0, final_scores=0.61207,
final_local_scores=0.644643, pem=-0.27389407,
similarity=24.816624]
67: 815988 William_Tanui
    [pem_pos=0, final_scores=0.46885282,
final_local_scores=0.60244286, pem=0.0, similarity=22.06781]
68: 188171 Kenya
    [pem_pos=0, final_scores=0.6118395,
final_local_scores=0.6443911, pem=-0.27389407,
similarity=24.809107]
69: 13182495 John_Mayock
    [pem_pos=0, final_scores=0.51611096,
final_local_scores=0.6523138, pem=0.0, similarity=23.556059]
70: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.58957344,
final_local_scores=0.56091636, pem=-0.61114657,
similarity=24.152086]
71: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.603865,
final_local_scores=0.5765278, pem=-0.61114657,
similarity=24.61796]
72: 805864 Roger_Black
    [pem_pos=0, final_scores=0.63136464,
final_local_scores=0.7260758, pem=0.0, similarity=25.757263]
```

```
73: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.6036585,
final_local_scores=0.5763021, pem=-0.61114657,
similarity=24.611229]
74: 7652893 Mark_Richardson_(athlete)
    [pem_pos=2, final_scores=0.52885157,
final_local_scores=0.6346206, pem=-1.0948459,
similarity=28.982]
75: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.61442775,
final_local_scores=0.5880659, pem=-0.61114657,
similarity=24.962284]
76: 1676095 Derek_Mills
    [pem_pos=0, final_scores=0.5192232,
final_local_scores=0.65230304, pem=0.0, similarity=23.555737]
77: 3434750 United_States
    [pem_pos=0, final_scores=0.48546022,
final_local_scores=0.45602936, pem=-0.14075686,
similarity=18.463997]
78: 2736304 Osmond_Ezinwa
    [pem_pos=0, final_scores=0.59925807,
final_local_scores=0.7516886, pem=0.0, similarity=26.5216]
79: 21383 Nigeria
    [pem_pos=0, final_scores=0.6345192,
final_local_scores=0.6651378, pem=-0.31455088,
similarity=25.64933]
80: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5825102,
final_local_scores=0.5532008, pem=-0.61114657,
similarity=23.921839]
81: 67118 Linford_Christie
    [pem_pos=0, final_scores=0.98635465,
final_local_scores=1.0935521, pem=0.0, similarity=36.723488]
82: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.58445823,
final_local_scores=0.5553288, pem=-0.61114657,
similarity=23.98534]
```

Sample: b'318_MOTOR'
MOTOR RACING - 1 BELGIAN GRAND PRIX RESULT . [1 1 SPA-
FRANCORCHAMPS 1] , [2 2 Belgium 2] 1996-08-25 Result of
Sunday 's 2 Belgian Grand Prix motor race : 1. [3 3 Michael Sc
humacher 3] ([4 4 Germany 4]) [5 5 Ferrari 5] 1 hour 28
minutes 15.125 seconds (average speed 208.442 kph) 2. [6 6 J
acques Villeneuve 6] ([7 7 Canada 7]) [8 8 Williams 8]
5.602 seconds behind 3. [9 9 Mika Hakkinen 9] ([10 10 Finland
10]) [11 11 McLaren 11] 15.710 4. [12 12 Jean Alesi 12] ([13
13 France 13]) [14 14 Benetton 14] 19.125 5. [15 15 Damon Hi
ll 15] ([16 16 Britain 16]) [17 17 Williams 17] 29.179 6. [1
8 18 Gerhard Berger 18] ([19 19 Austria 19]) [20 20 Benetton
20] 29.896 7. [21 21 Mika Salo 21] ([22 22 Finland 22]) [23
23 Tyrrell 23] 1:00.754 8. [24 24 Ukyo Katayama 24] ([25 25 J

apan 25]) [26 26 Tyrrell 26] 1:40.227 9. [27 27 Ricardo Ross et 27] ([28 28 Brazil 28]) [29 29 Arrows 29] one lap 10. [30 30 Pedro Lamy 30] ([31 31 Portugal 31]) [32 32 Minardi 32] one lap Did not finish : 11. [33 33 David Coulthard 33] ([34 34 Britain 34]) [35 35 McLaren 35] 37 laps completed 12. [36 36 Martin Brundle 36] ([37 37 Britain 37]) [38 38 Jordan 38] 34 13. [39 39 Eddie Irvine 39] ([40 40 Britain 40]) [41 41 F errari 41] 29 14. [42 42 Rubens Barrichello 42] ([43 43 Braz il 43]) [44 44 Jordan 44] 29 15. [45 45 Pedro Diniz 45] ([46 46 Brazil 46]) [47 47 Ligier 47] 22 16. [48 48 Jos Verstappen 48] ([49 49 Netherland 49]) [50 50 Arrows 50] 11 Did not start (failed to complete one lap) : [51 51 Olivier Panis 51] ([52 52 France 52]) [53 53 Ligier 53] [54 54 Johnny Herbe rt 54] ([55 55 Britain 55]) [56 56 Sauber 56] [57 57 Heinz-Harald Frentzen 57] ([58 58 Germany 58]) [59 59 Sauber 59] Fastest lap : [60 60 Berger 60] 1:53.067 (221.857 kph)
1: 948477 Circuit_de_Spa-Francorchamps
2: 3343 Belgium
3: 20396 Michael_Schumacher
4: 11867 Germany
5: 173009 Scuderia_Ferrari
6: 64258 Jacques_Villeneuve
7: 5042916 Canada
8: 34104 Williams_F1
9: 63073 Mika_Häkkinen
10: 10577 Finland
11: 20994 McLaren
12: 341553 Jean_Alesi
13: 5843419 France
14: 445741 Benetton_Formula
15: 78249 Damon_Hill
16: 31717 United_Kingdom
17: 34104 Williams_F1
18: 343833 Gerhard_Berger
19: 26964606 Austria
20: 445741 Benetton_Formula
21: 1157079 Mika_Salo
22: 10577 Finland
23: 342015 Tyrrell_Racing
24: 520249 Ukyo_Katayama
25: 15573 Japan
26: 342015 Tyrrell_Racing
27: 1226544 Ricardo_Rosset
28: 3383 Brazil
29: 9573077 Footwork_Arrows
30: 909656 Pedro_Lamy
31: 23033 Portugal
32: 40479 Minardi
33: 19636733 David_Coulthard
34: 31717 United_Kingdom
35: 20994 McLaren
36: 341560 Martin_Brundle

```
37: 31717 United_Kingdom
38: 458472 Jordan_Grand_Prix
39: 424417 Eddie_Irvine
40: 31717 United_Kingdom
41: 173009 Scuderia_Ferrari
42: 252989 Rubens_Barrichello
43: 3383 Brazil
44: 458472 Jordan_Grand_Prix
45: 1226332 Pedro_Diniz
46: 3383 Brazil
47: 24726570 Equipe_Ligier
48: 1148735 Jos_Verstappen
49: 21148 Netherlands
50: 9573077 Footwork_Arrows
51: 414513 Olivier_Panis
52: 5843419 France
53: 24726570 Equipe_Ligier
54: 342991 Johnny_Herbert
55: 31717 United_Kingdom
56: 3158734 Sauber
57: 756775 Heinz-Harald_Frentzen
58: 11867 Germany
59: 3158734 Sauber
60: 343833 Gerhard_Berger
1: 1123283 1996_Belgian_Grand_Prix
    [winner: 679716 Belgium_national_football_team,
pem_pos=18, final_scores=-0.108419105,
final_local_scores=-0.07902467, pem=-3.9058492,
similarity=22.972075]
    RECALL MISS: 1123283 1996_Belgian_Grand_Prix
    [winner: 54221 Grand_Prix_motorcycle_racing, pem_pos=1,
final_scores=-0.21255752, final_local_scores=-0.26329172,
pem=-1.5994394, similarity=4.9305625]
    RECALL MISS: 1123283 1996_Belgian_Grand_Prix]
2: 1123283 1996_Belgian_Grand_Prix
    [winner: 351087 Belgian_Grand_Prix, pem_pos=0,
final_scores=-0.07293208, final_local_scores=-0.11810196,
pem=-0.50572526, similarity=3.3155274]
    RECALL MISS: 1123283 1996_Belgian_Grand_Prix]
1: 948477 Circuit_de_Spa-Francorchamps
    [pem_pos=0, final_scores=0.38668185,
final_local_scores=0.36325824, pem=0.0, similarity=14.930063]
2: 3343 Belgium
    [pem_pos=0, final_scores=0.53313196,
final_local_scores=0.58460253, pem=-0.24180195,
similarity=22.850376]
3: 20396 Michael_Schumacher
    [pem_pos=0, final_scores=0.56970924,
final_local_scores=0.5462402, pem=0.0, similarity=20.39061]
4: 11867 Germany
    [pem_pos=0, final_scores=0.7072181,
final_local_scores=0.793461, pem=-0.15529114,
```

```
similarity=28.61267]
5: 173009 Scuderia_Ferrari
    [pem_pos=1, final_scores=0.42595667,
final_local_scores=0.385153, pem=-0.77418387,
similarity=19.79358]
6: 64258 Jacques_Villeneuve
    [pem_pos=0, final_scores=0.43432635,
final_local_scores=0.40694094, pem=0.0, similarity=16.233643]
7: 5042916 Canada
    [pem_pos=0, final_scores=0.7069703,
final_local_scores=0.8055146, pem=-0.11662793,
similarity=28.762115]
8: 34104 Williams_F1
    [pem_pos=0, final_scores=0.66947573,
final_local_scores=0.657353, pem=-1.0414248,
similarity=29.369867]
9: 63073 Mika_Häkkinen
    [pem_pos=0, final_scores=0.80699056,
final_local_scores=0.81576073, pem=0.0, similarity=28.43364]
10: 10577 Finland
    [pem_pos=0, final_scores=0.7382391,
final_local_scores=0.83556306, pem=-0.2445894,
similarity=30.354698]
11: 20994 McLaren
    [pem_pos=0, final_scores=0.47943014,
final_local_scores=0.4408257, pem=-0.11262372,
similarity=17.857296]
12: 341553 Jean_Alesi
    [pem_pos=0, final_scores=0.5412861,
final_local_scores=0.52346784, pem=0.0, similarity=19.711037]
13: 5843419 France
    [pem_pos=0, final_scores=0.6764575,
final_local_scores=0.7468151, pem=-0.16240834,
similarity=27.259365]
14: 445741 Benetton_Formula
    [pem_pos=0, final_scores=0.3707686,
final_local_scores=0.3401479, pem=-0.069357015,
similarity=14.617579]
15: 78249 Damon_Hill
    [pem_pos=0, final_scores=0.6104526,
final_local_scores=0.5980197, pem=0.0, similarity=21.935814]
16: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.4981801,
final_local_scores=0.5634417, pem=-0.61114657,
similarity=24.227448]
17: 34104 Williams_F1
    [pem_pos=0, final_scores=0.5865766,
final_local_scores=0.5667982, pem=-1.0414248,
similarity=26.667532]
18: 343833 Gerhard_Berger
    [pem_pos=0, final_scores=0.5909404,
final_local_scores=0.58217937, pem=0.0, similarity=21.463106]
```

```
19: 26964606 Austria
    [pem_pos=0, final_scores=0.6201317,
final_local_scores=0.69758624, pem=-0.25262418,
similarity=26.280888]
20: 445741 Benetton_Formula
    [pem_pos=0, final_scores=0.34899473,
final_local_scores=0.31636328, pem=-0.069357015,
similarity=13.907799]
21: 1157079 Mika_Salo
    [pem_pos=0, final_scores=0.7032569,
final_local_scores=0.7205646, pem=0.0, similarity=25.592798]
22: 10577 Finland
    [pem_pos=0, final_scores=0.70906657,
final_local_scores=0.8036965, pem=-0.2445894,
similarity=29.403736]
23: 342015 Tyrrell_Racing
    [pem_pos=0, final_scores=0.4799462,
final_local_scores=0.4577728, pem=-0.054408323,
similarity=18.046448]
24: 520249 Ukyo_Katayama
    [pem_pos=0, final_scores=0.52736324,
final_local_scores=0.5556001, pem=0.0, similarity=20.669928]
25: 15573 Japan
    [pem_pos=0, final_scores=0.6796039,
final_local_scores=0.78450847, pem=-0.13549899,
similarity=28.237875]
26: 342015 Tyrrell_Racing
    [pem_pos=0, final_scores=0.515907,
final_local_scores=0.49705458, pem=-0.054408323,
similarity=19.218695]
27: 1226544 Ricardo_Rosset
    [pem_pos=0, final_scores=0.6330556,
final_local_scores=0.69706583, pem=0.0, similarity=24.891548]
28: 3383 Brazil
    [pem_pos=0, final_scores=0.6913161,
final_local_scores=0.79259676, pem=-0.21313342,
similarity=28.901434]
29: 9573077 Footwork_Arrows
    [pem_pos=5, final_scores=0.3330875,
final_local_scores=0.32401633, pem=-2.0794415,
similarity=25.067333]
30: 909656 Pedro_Lamy
    [pem_pos=0, final_scores=0.4648311,
final_local_scores=0.48198444, pem=0.0, similarity=18.47309]
31: 23033 Portugal
    [pem_pos=0, final_scores=0.5701718,
final_local_scores=0.6524032, pem=-0.30942732,
similarity=25.241438]
32: 40479 Minardi
    [pem_pos=0, final_scores=0.3713019,
final_local_scores=0.34838536, pem=0.0, similarity=14.486227]
33: 19636733 David_Coulthard
```

```
[pem_pos=0, final_scores=0.6608094,
final_local_scores=0.6533075, pem=0.0, similarity=23.585712]
34: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.50003755,
final_local_scores=0.56547064, pem=-0.61114657,
similarity=24.287994]
35: 20994 McLaren
    [pem_pos=0, final_scores=0.48938632,
final_local_scores=0.45170134, pem=-0.11262372,
similarity=18.181847]
36: 341560 Martin_Brundle
    [pem_pos=0, final_scores=0.6501527,
final_local_scores=0.66041493, pem=0.0, similarity=23.797812]
37: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.54286945,
final_local_scores=0.612258, pem=-0.61114657,
similarity=25.684223]
38: 458472 Jordan_Grand_Prix
    [pem_pos=1, final_scores=0.7981369,
final_local_scores=0.8009974, pem=-2.8876953,
similarity=43.696808]
39: 424417 Eddie_Irvine
    [pem_pos=0, final_scores=0.4669724,
final_local_scores=0.4390865, pem=0.0, similarity=17.19293]
40: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5515606,
final_local_scores=0.6217517, pem=-0.61114657,
similarity=25.967533]
41: 173009 Scuderia_Ferrari
    [pem_pos=1, final_scores=0.28633362,
final_local_scores=0.23263597, pem=-0.77418387,
similarity=15.242167]
42: 252989 Rubens_Barrichello
    [pem_pos=0, final_scores=0.54353124,
final_local_scores=0.52710605, pem=0.0, similarity=19.819609]
43: 3383 Brazil
    [pem_pos=0, final_scores=0.61328465,
final_local_scores=0.70735925, pem=-0.21313342,
similarity=26.357777]
44: 458472 Jordan_Grand_Prix
    [pem_pos=1, final_scores=0.7302241,
final_local_scores=0.72681296, pem=-2.8876953,
similarity=41.482998]
45: 1226332 Pedro_Diniz
    [pem_pos=0, final_scores=0.51064914,
final_local_scores=0.5326944, pem=0.0, similarity=19.986376]
46: 3383 Brazil
    [pem_pos=0, final_scores=0.63362855,
final_local_scores=0.7295819, pem=-0.21313342,
similarity=27.020945]
47: 24726570 Equipe_Ligier
    [pem_pos=1, final_scores=0.37756568,
```

```
final_local_scores=0.35227174, pem=-0.5983149,
similarity=17.855934]
48: 1148735 Jos_Verstappen
    [pem_pos=0, final_scores=0.49747306,
final_local_scores=0.4818611, pem=0.0, similarity=18.46941]
49: 21148 Netherlands
    [pem_pos=1, final_scores=0.262307,
final_local_scores=0.30418783, pem=-1.0152442,
similarity=18.68834]
50: 9573077 Footwork_Arrows
    [pem_pos=5, final_scores=0.40036207,
final_local_scores=0.39750355, pem=-2.0794415,
similarity=27.260338]
51: 414513 Olivier_Panis
    [pem_pos=0, final_scores=0.54467684,
final_local_scores=0.548593, pem=0.0, similarity=20.460821]
52: 5843419 France
    [pem_pos=0, final_scores=0.6488031,
final_local_scores=0.71660686, pem=-0.16240834,
similarity=26.357891]
53: 24726570 Equipe_Ligier
    [pem_pos=1, final_scores=0.41499776,
final_local_scores=0.39316058, pem=-0.5983149,
similarity=19.07614]
54: 342991 Johnny_Herbert
    [pem_pos=0, final_scores=0.71396357,
final_local_scores=0.71692365, pem=0.0, similarity=25.484144]
55: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5267138,
final_local_scores=0.5946104, pem=-0.61114657,
similarity=25.157583]
56: 3158734 Sauber
    [pem_pos=0, final_scores=0.3132835,
final_local_scores=0.27606648, pem=0.0, similarity=12.328088]
57: 756775 Heinz-Harald_Frentzen
    [pem_pos=0, final_scores=0.71098083,
final_local_scores=0.714499, pem=0.0, similarity=25.411789]
58: 11867 Germany
    [pem_pos=0, final_scores=0.75762445,
final_local_scores=0.84852225, pem=-0.15529114,
similarity=30.255808]
59: 3158734 Sauber
    [pem_pos=0, final_scores=0.3998589,
final_local_scores=0.37063694, pem=0.0, similarity=15.150259]
60: 343833 Gerhard_Berger
    [pem_pos=0, final_scores=1.2168593,
final_local_scores=1.2685513, pem=0.0, similarity=41.945812]
```

Sample: b'319_MOTOR'

MOTOR RACING - [1 1 SCHUMACHER 1] WINS 1 BELGIAN GRAND
PRIX . SPA-FRANCORCHAMPS 1996-08-25 [2 2 Michael Schumacher 2]
of [3 3 Germany 3] , driving a [4 4 Ferrari 4] , won the 2 Be

lgian Grand Prix motor race on Sunday . [5 5 Canada 5] 's [6 6 Jacques Villeneuve 6] finished second in his [7 7 Williams 7] and [8 8 Mika Hakkinen 8] of [9 9 Finland 9] was third in a [1 0 10 McLaren 10] . [11 11 Frenchman 11] [12 12 Jean Alesi 12] came fourth in his [13 13 Benetton 13] with [14 14 Britain 14] 's [15 15 Damon Hill 15] fifth in a [16 16 Williams 16] and [1 7 17 Gerhard Berger 17] of [18 18 Austria 18] sixth in the other [19 19 Benetton 19] . World drivers ' championship standings (after 13 rounds) : 1. [20 20 Damon Hill 20] ([21 21 Britain 21]) 81 points 2. [22 22 Jacques Villeneuve 22] ([23 23 Canada 23]) 68 3. [24 24 Michael Schumacher 24] ([25 25 Germany 25]) 39 4. [26 26 Jean Alesi 26] ([27 27 France 2 7]) 38 5. [28 28 Mika Hakkinen 28] ([29 29 Finland 29]) 23 6. [30 30 David Coulthard 30] ([31 31 Britain 31]) 18 7. [32 32 Gerhard Berger 32] ([33 33 Austria 33]) 17 8. [34 34 Olivier Panis 34] ([35 35 France 35]) 13 9. [36 36 Rubens Barrichello 36] ([37 37 Brazil 37]) 12 10. [38 38 Eddie Irvine 3 8] ([39 39 Britain 39]) 9 11. [40 40 Heinz-Harald Frentzen 4 0] ([41 41 Germany 41]) 6 12. [42 42 Mika Salo 42] ([43 43 Finland 43]) 5 13. [44 44 Johnny Herbert 44] ([45 45 Britain 45]) 4 14. [46 46 Martin Brundle 46] ([47 47 Britain 47]) 3 15 equal . [48 48 Jos Verstappen 48] ([49 49 Netherlands 49]) 1 15 equal . [50 50 Pedro Diniz 50] ([51 51 Brazil 51]) 1 Constructors ' championship : 1. [52 52 Williams 52] 149 points 2. [53 53 Benetton 53] 55 3. [54 54 Ferrari 54] 48 4. [55 55 McLaren 55] 41 5. [56 56 Jordan 56] 15 6. [57 57 Ligier 57] 14 7. [58 58 Sauber 58] 10 8. [59 59 Tyrrell 59] 5 9.

Footwork 1

1: 20396 Michael_Schumacher
2: 20396 Michael_Schumacher
3: 11867 Germany
4: 11225 Ferrari
5: 5042916 Canada
6: 64258 Jacques_Villeneuve
7: 34104 Williams_F1
8: 63073 Mika_Häkkinen
9: 10577 Finland
10: 20994 McLaren
11: 5843419 France
12: 341553 Jean_Alesi
13: 445741 Benetton_Formula
14: 31717 United_Kingdom
15: 78249 Damon_Hill
16: 34104 Williams_F1
17: 343833 Gerhard_Berger
18: 26964606 Austria
19: 445741 Benetton_Formula
20: 78249 Damon_Hill
21: 31717 United_Kingdom
22: 64258 Jacques_Villeneuve
23: 5042916 Canada
24: 20396 Michael_Schumacher

```
25: 11867 Germany
26: 341553 Jean_Alesi
27: 5843419 France
28: 63073 Mika_Häkkinen
29: 10577 Finland
30: 19636733 David_Coulthard
31: 31717 United_Kingdom
32: 343833 Gerhard_Berger
33: 26964606 Austria
34: 414513 Olivier_Panis
35: 5843419 France
36: 252989 Rubens_Barrichello
37: 3383 Brazil
38: 424417 Eddie_Irvine
39: 31717 United_Kingdom
40: 756775 Heinz-Harald_Frentzen
41: 11867 Germany
42: 1157079 Mika_Salo
43: 10577 Finland
44: 342991 Johnny_Herbert
45: 31717 United_Kingdom
46: 341560 Martin_Brundle
47: 31717 United_Kingdom
48: 1148735 Jos_Verstappen
49: 21148 Netherlands
50: 1226332 Pedro_Diniz
51: 3383 Brazil
52: 34104 Williams_F1
53: 445741 Benetton_Formula
54: 11225 Ferrari
55: 20994 McLaren
56: 458472 Jordan_Grand_Prix
57: 24726570 Equipe_Ligier
58: 3158734 Sauber
59: 342015 Tyrrell_Racing
1: 1123283 1996_Belgian_Grand_Prix
    [winner: 3343 Belgium, pem_pos=0,
final_scores=-0.18668742, final_local_scores=-0.21715428,
pem=-0.12403572, similarity=-1.7160783]
    RECALL MISS: 1123283 1996_Belgian_Grand_Prix]
    [winner: 54221 Grand_Prix_motorcycle_racing, pem_pos=1,
final_scores=-0.30047864, final_local_scores=-0.35522377,
pem=-1.5994394, similarity=2.1871266|
    RECALL MISS: 1123283 1996_Belgian_Grand_Prix]
2: 1123283 1996_Belgian_Grand_Prix
    [winner: 351087 Belgian_Grand_Prix, pem_pos=0,
final_scores=-0.080847695, final_local_scores=-0.11988255,
pem=-0.50572526, similarity=3.2623913|
    RECALL MISS: 1123283 1996_Belgian_Grand_Prix]
1: 20396 Michael_Schumacher
    [pem_pos=0, final_scores=0.58398026,
final_local_scores=0.5640883, pem=0.0, similarity=20.923233]
```

```
2: 20396 Michael_Schumacher
    [pem_pos=0, final_scores=0.4943623,
final_local_scores=0.4640395, pem=0.0, similarity=17.937578]
3: 11867 Germany
    [pem_pos=0, final_scores=0.7071072,
final_local_scores=0.7776944, pem=-0.15529114,
similarity=28.142162]
4: 11225 Ferrari
    [pem_pos=0, final_scores=0.281435,
final_local_scores=0.26298696, pem=-0.64550596,
similarity=15.448131]
5: 5042916 Canada
    [pem_pos=0, final_scores=0.671506,
final_local_scores=0.75137484, pem=-0.11662793,
similarity=27.146477]
6: 64258 Jacques_Villeneuve
    [pem_pos=0, final_scores=0.36670202,
final_local_scores=0.33484113, pem=0.0, similarity=14.08204]
7: 34104 Williams_F1
    [pem_pos=1, final_scores=0.5613488,
final_local_scores=0.5453086, pem=-1.0414248,
similarity=26.02624]
8: 63073 Mika_Häkkinen
    [pem_pos=0, final_scores=0.72467524,
final_local_scores=0.7271389, pem=0.0, similarity=25.788988]
9: 10577 Finland
    [pem_pos=0, final_scores=0.69112176,
final_local_scores=0.7679173, pem=-0.2445894,
similarity=28.336012]
10: 20994 McLaren
    [pem_pos=0, final_scores=0.3893301,
final_local_scores=0.34866005, pem=-0.11262372,
similarity=15.10689]
11: 5843419 France
    [pem_pos=1, final_scores=0.41937733,
final_local_scores=0.45427543, pem=-0.7184992,
similarity=21.553509]
12: 341553 Jean_Alesi
    [pem_pos=0, final_scores=0.4644276,
final_local_scores=0.44655037, pem=0.0, similarity=17.415667]
13: 445741 Benetton_Formula
    [pem_pos=0, final_scores=0.3045929,
final_local_scores=0.27375293, pem=-0.069357015,
similarity=12.636221]
14: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.4907034,
final_local_scores=0.5390592, pem=-0.61114657,
similarity=23.499826]
15: 78249 Damon_Hill
    [pem_pos=0, final_scores=0.5371284,
final_local_scores=0.5187579, pem=0.0, similarity=19.570482]
16: 34104 Williams_F1
```

```
[pem_pos=0, final_scores=0.55264145,
final_local_scores=0.5357972, pem=-1.0414248,
similarity=25.7424]
17: 343833 Gerhard_Berger
    [pem_pos=0, final_scores=0.52394557,
final_local_scores=0.51932204, pem=0.0, similarity=19.587318]
18: 26964606 Austria
    [pem_pos=0, final_scores=0.38257664,
final_local_scores=0.42191327, pem=-0.25262418,
similarity=18.054256]
19: 445741 Benetton_Formula
    [pem_pos=0, final_scores=0.31855148,
final_local_scores=0.28900057, pem=-0.069357015,
similarity=13.091241]
20: 78249 Damon_Hill
    [pem_pos=0, final_scores=0.6141258,
final_local_scores=0.6028658, pem=0.0, similarity=22.080431]
21: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.48814613,
final_local_scores=0.5362658, pem=-0.61114657,
similarity=23.416464]
22: 64258 Jacques_Villeneuve
    [pem_pos=0, final_scores=0.40625268,
final_local_scores=0.3780443, pem=0.0, similarity=15.371308]
23: 5042916 Canada
    [pem_pos=0, final_scores=0.6880525,
final_local_scores=0.76944935, pem=-0.11662793,
similarity=27.685858]
24: 20396 Michael_Schumacher
    [pem_pos=0, final_scores=0.6403418,
final_local_scores=0.62349993, pem=0.0, similarity=22.696194]
25: 11867 Germany
    [pem_pos=0, final_scores=0.76235265,
final_local_scores=0.8380417, pem=-0.15529114,
similarity=29.943047]
26: 341553 Jean_Alesi
    [pem_pos=0, final_scores=0.5849301,
final_local_scores=0.57818115, pem=0.0, similarity=21.343792]
27: 5843419 France
    [pem_pos=0, final_scores=0.70540494,
final_local_scores=0.7667172, pem=-0.16240834,
similarity=27.853285]
28: 63073 Mika_Häkkinen
    [pem_pos=0, final_scores=0.7924909,
final_local_scores=0.8012172, pem=0.0, similarity=27.999632]
29: 10577 Finland
    [pem_pos=0, final_scores=0.7448067,
final_local_scores=0.8265599, pem=-0.2445894,
similarity=30.086025]
30: 19636733 David_Coulthard
    [pem_pos=0, final_scores=0.69925404,
final_local_scores=0.69736725, pem=0.0, similarity=24.900543]
```

```
31: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5394297,
final_local_scores=0.59228534, pem=-0.61114657,
similarity=25.088198]
32: 343833 Gerhard_Berger
    [pem_pos=0, final_scores=0.63697416,
final_local_scores=0.6427886, pem=0.0, similarity=23.271807]
33: 26964606 Austria
    [pem_pos=0, final_scores=0.5668181,
final_local_scores=0.62316906, pem=-0.25262418,
similarity=24.060131]
34: 414513 Olivier_Panis
    [pem_pos=0, final_scores=0.5691816,
final_local_scores=0.5826575, pem=0.0, similarity=21.477375]
35: 5843419 France
    [pem_pos=0, final_scores=0.6585308,
final_local_scores=0.7155143, pem=-0.16240834,
similarity=26.325289]
36: 252989 Rubens_Barrichello
    [pem_pos=0, final_scores=0.540623,
final_local_scores=0.5282892, pem=0.0, similarity=19.854916]
37: 3383 Brazil
    [pem_pos=0, final_scores=0.72040766,
final_local_scores=0.81857, pem=-0.21313342,
similarity=29.676527]
38: 424417 Eddie_Irvine
    [pem_pos=0, final_scores=0.5103918,
final_local_scores=0.48973304, pem=0.0, similarity=18.704323]
39: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5281876,
final_local_scores=0.580005, pem=-0.61114657,
similarity=24.72173]
40: 756775 Heinz-Harald_Frentzen
    [pem_pos=0, final_scores=0.8802126,
final_local_scores=0.9080996, pem=0.0, similarity=31.189217]
41: 11867 Germany
    [pem_pos=0, final_scores=0.73530865,
final_local_scores=0.80850023, pem=-0.15529114,
similarity=29.06147]
42: 1157079 Mika_Salo
    [pem_pos=0, final_scores=0.6439226,
final_local_scores=0.66581637, pem=0.0, similarity=23.959002]
43: 10577 Finland
    [pem_pos=0, final_scores=0.70189923,
final_local_scores=0.77969, pem=-0.2445894,
similarity=28.687334]
44: 342991 Johnny_Herbert
    [pem_pos=0, final_scores=0.69885176,
final_local_scores=0.7093249, pem=0.0, similarity=25.257383]
45: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.54504734,
final_local_scores=0.59842175, pem=-0.61114657,
```

```
similarity=25.27132]
46: 341560 Martin_Brundle
    [pem_pos=0, final_scores=0.6171858,
final_local_scores=0.6314251, pem=0.0, similarity=22.932697]
47: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5105871,
final_local_scores=0.56077915, pem=-0.61114657,
similarity=24.147991]
48: 1148735 Jos_Verstappen
    [pem_pos=0, final_scores=0.49524313,
final_local_scores=0.49088985, pem=0.0, similarity=18.738846]
49: 21148 Netherlands
    [pem_pos=0, final_scores=0.55118877,
final_local_scores=0.6085767, pem=-0.24940795,
similarity=23.607176]
50: 1226332 Pedro_Diniz
    [pem_pos=0, final_scores=0.50037056,
final_local_scores=0.53098613, pem=0.0, similarity=19.935398]
51: 3383 Brazil
    [pem_pos=0, final_scores=0.65720034,
final_local_scores=0.7495256, pem=-0.21313342,
similarity=27.616104]
52: 34104 Williams_F1
    [pem_pos=1, final_scores=0.6321791,
final_local_scores=0.6226799, pem=-1.0414248,
similarity=28.335152]
53: 445741 Benetton_Formula
    [pem_pos=0, final_scores=0.40069503,
final_local_scores=0.37872994, pem=-0.069357015,
similarity=15.768944]
54: 11225 Ferrari
    [pem_pos=0, final_scores=0.2137308,
final_local_scores=0.1900889, pem=-0.64550596,
similarity=13.272707]
55: 20994 McLaren
    [pem_pos=0, final_scores=0.5132086,
final_local_scores=0.48397857, pem=-0.11262372,
similarity=19.145063]
56: 458472 Jordan_Grand_Prix
    [pem_pos=1, final_scores=0.6588499,
final_local_scores=0.65885746, pem=-2.8876953,
similarity=39.45507]
57: 24726570 Equipe_Ligier
    [pem_pos=1, final_scores=0.3898738,
final_local_scores=0.37848133, pem=-0.5983149,
similarity=18.63808]
58: 3158734 Sauber
    [pem_pos=0, final_scores=0.41327137,
final_local_scores=0.3945588, pem=0.0, similarity=15.864134]
59: 342015 Tyrrell_Racing
    [pem_pos=0, final_scores=0.5261107,
final_local_scores=0.519397, pem=-0.054408323,
```

```
similarity=19.885437]
```

```
Sample: b'320_RALLYING'
RALLYING - LEADING POSITIONS IN 1,000 LAKES RALLY . [1 1 JYV
ASKYLA 1] , [2 2 Finland 2] 1996-08-25 Leading positions
after six of Sunday 's 12 special stages in the 1,000 Lakes
Rally , sixth round of the world championship : 1. [3 3 Juha K
ankkunen 3] ( [4 4 Finland 4] ) [5 5 Toyota Celica 5] 2 hours
30 minutes 52 seconds 3. Tommi Makinen ( [6 6 Finland 6] ) [7
7 Mitsubishi Lancer 7] 8 seconds behind 2. [8 8 Marcus Grönho
lm 8] ( [9 9 Finland 9] ) [10 10 Toyota Celica 10] 1:46 4. [11
11 Jarmo Kytölehto 11] ( [12 12 Finland 12] ) Ford Escort
1:56 5. [13 13 Kenneth Eriksson 13] ( [14 14 Sweden 14] ) [15
15 Subaru Impreza 15] 2:05 6. [16 16 Thomas Radstrom 16] ( [17
17 Sweden 17] ) [18 18 Toyota Celica 18] 2:23
1: 258746 Jyväskylä
2: 10577 Finland
3: 412742 Juha_Kankunen
4: 10577 Finland
5: 172542 Toyota_Celica
6: 10577 Finland
7: 1119339 Mitsubishi_Lancer
8: 944600 Marcus_Grönholm
9: 10577 Finland
10: 172542 Toyota_Celica
11: 20730137 Jarmo_Kytölehto
12: 10577 Finland
13: 1411702 Kenneth_Eriksson
14: 5058739 Sweden
15: 527976 Subaru_Impreza
16: 7364013 Thomas_Rådström
17: 5058739 Sweden
18: 172542 Toyota_Celica
1: 258746 Jyväskylä
    [pem_pos=0, final_scores=0.50710267,
final_local_scores=0.560583, pem=0.0, similarity=20.818628]
2: 10577 Finland
    [pem_pos=0, final_scores=0.799784,
final_local_scores=0.8147795, pem=-0.2445894,
similarity=29.734474]
3: 412742 Juha_Kankunen
    [pem_pos=0, final_scores=0.5888,
final_local_scores=0.63977736, pem=0.0, similarity=23.181946]
4: 10577 Finland
    [pem_pos=0, final_scores=0.7392009,
final_local_scores=0.7486016, pem=-0.2445894,
similarity=27.759594]
5: 172542 Toyota_Celica
    [pem_pos=0, final_scores=0.39020193,
final_local_scores=0.4072166, pem=0.0, similarity=16.241869]
6: 10577 Finland
    [pem_pos=0, final_scores=0.7726479,
```

```
final_local_scores=0.7851374, pem=-0.2445894,
similarity=28.849895]
7: 1119339 Mitsubishi_Lancer
    [pem_pos=0, final_scores=0.29905355,
final_local_scores=0.35447457, pem=0.0, similarity=14.667941]
8: 944600 Marcus_Grönholm
    [pem_pos=0, final_scores=0.8263226,
final_local_scores=0.8810422, pem=0.0, similarity=30.381767]
9: 10577 Finland
    [pem_pos=0, final_scores=0.77711314,
final_local_scores=0.790015, pem=-0.2445894,
similarity=28.995451]
10: 172542 Toyota_Celica
    [pem_pos=0, final_scores=0.45009553,
final_local_scores=0.4726413, pem=0.0, similarity=18.194271]
11: 20730137 Jarmo_Kytölehto
    [pem_pos=0, final_scores=0.511409,
final_local_scores=0.6372997, pem=0.0, similarity=23.108007]
12: 10577 Finland
    [pem_pos=0, final_scores=0.71521765,
final_local_scores=0.7224036, pem=-0.2445894,
similarity=26.977793]
13: 1411702 Kenneth_Eriksson
    [pem_pos=0, final_scores=0.5485146,
final_local_scores=0.6201657, pem=0.0, similarity=22.596695]
14: 5058739 Sweden
    [pem_pos=0, final_scores=0.76183814,
final_local_scores=0.7941904, pem=-0.2135508,
similarity=28.951262]
15: 527976 Subaru_Impreza
    [pem_pos=0, final_scores=0.14725463,
final_local_scores=0.14810467, pem=0.0, similarity=8.509452]
16: 7364013 Thomas_Rådström
    [pem_pos=0, final_scores=0.4874388,
final_local_scores=0.5720634, pem=0.0, similarity=21.161226]
17: 5058739 Sweden
    [pem_pos=0, final_scores=0.7661075,
final_local_scores=0.798854, pem=-0.2135508,
similarity=29.090431]
18: 172542 Toyota_Celica
    [pem_pos=0, final_scores=0.3887859,
final_local_scores=0.4056698, pem=0.0, similarity=16.19571]
```

Sample: b'321_MOTORCYCLING'

MOTORCYCLING - WORLD SUPERBIKE CHAMPIONSHIP RESULTS . SUGO ,
[1 1 Japan 1] 1996-08-25 Leading results from round nine of
the superbike world championship on Sunday : First race 1.
Yuuchi Takeda ([2 2 Japan 2]) [3 3 Honda 3] 38 minutes
30.054 seconds 2. [4 4 Noriyuki Haga 4] ([5 5 Japan 5]) [6 6
Yamaha 6] 38:30.140 3. [7 7 Wataru Yoshikawa 7] ([8 8 Japan 8
]) [9 9 Yamaha 9] 38:32.353 4. [10 10 Troy Corser 10] ([11 1
1 Australia 11]) [12 12 Ducati 12] 38:34.436 5. [13 13 John K

ocinski 13] ([14 14 U.S. 14]) [15 15 Ducati 15] 38:36.306
6. [16 16 Aaron Slight 16] ([17 17 New Zealand 17]) [18 18 H
onda 18] 38:41.756 7. Norihiko Fujiwara ([19 19 Japan 19]) [
20 20 Yamaha 20] 38:43.253 8. [21 21 Carl Fogarty 21] ([22 22
Britain 22]) [23 23 Honda 23] 38:49.595 9. [24 24 Akira Ryo 2
4] ([25 25 Japan 25]) [26 26 Kawasaki 26] 38:50.269 10.
Shiya Takeishi ([27 27 Japan 27]) [28 28 Kawasaki 28]
38:52.271 Fastest lap : [29 29 Haga 29] 147.159 kph . Second
race 1. [30 30 Takuma Aoki 30] ([31 31 Japan 31]) [32 32 Ho
nda 32] 38:18.759 2. [33 33 Kocinski 33] 38:19.313 3. [34 34 H
aga 34] 38:32.040 4. [35 35 Slight 35] 38:32.149 5. [36 36 Fo
garty 36] 38:32.719 6. Fujiwara 38:33.595 7. Ryo 38:34.682 8.
Takeishi 38:34.999 9. [37 37 Yoshikawa 37] 38:35.297 10. [38 3
8 Corser 38] 38:42.015 Fastest lap : [39 39 Aoki 39] 147.786
kph World championship standings (after nine rounds) : 1. [4
0 40 Slight 40] 280 points 2. [41 41 Corser 41] 269 3. [42 42
Kocinski 42] 254 4. [43 43 Fogarty 43] 236 5. [44 44 Colin Ed
wards 44] ([45 45 U.S. 45]) [46 46 Yamaha 46] 176 6. Pier
Francesco Chili ([47 47 Italy 47]) [48 48 Ducati 48] 175 7.
[49 49 Simon Crafar 49] ([50 50 New Zealand 50]) [51 51 Kaw
asaki 51] 132 8. [52 52 Anthony Gobert 52] ([53 53 Australia
53]) [54 54 Kawasaki 54] 117 9. [55 55 Yoshikawa 55] 107 10.
[56 56 Neil Hodgson 56] ([57 57 Britain 57]) [58 58 Ducati 5
8] 82 Revised placings for second race after the
disqualification of [59 59 Japanese 59] rider [60 60 Noriyuki
Haga 60] for using an illegal carburettor part : 1. [61 61 Ta
kuma Aoki 61] ([62 62 Japan 62]) [63 63 Honda 63] 38:18.759
2. [64 64 Kocinski 64] 38:19.313 3. [65 65 Slight 65]
38:32.149 4. [66 66 Fogarty 66] 38:32.719 5. Fujiwara
38:33.595 6. Ryo 38:34.682 7. Takeishi 38:34.999 8. [67 67 Yo
shikawa 67] 38:35.297 9. [68 68 Corser 68] 38:42.015 10.
Keiichi Kitigawa ([69 69 Japan 69]) [70 70 Suzuki 70]
38:42.333 Fastest lap : [71 71 Aoki 71] 147.786 kph Revised
world championship standings (after nine rounds) : 1. [72 72
Slight 72] 283 points 2. [73 73 Corser 73] 270 3. [74 74 Koci
nski 74] 254 4. [75 75 Fogarty 75] 238 5. [76 76 Colin Edwards
76] ([77 77 U.S. 77]) [78 78 Yamaha 78] 176 6. Pier
Francesco Chili ([79 79 Italy 79]) [80 80 Ducati 80] 175 7.
[81 81 Simon Crafar 81] ([82 82 New Zealand 82]) [83 83 Kaw
asaki 83] 133 8. [84 84 Anthony Gobert 84] ([85 85 Australia
85]) [86 86 Kawasaki 86] 117 9. [87 87 Yoshikawa 87] 108 10.
[88 88 Neil Hodgson 88] ([89 89 Britain 89]) [90 90 Ducati 9
0] 82
1: 15573 Japan
2: 15573 Japan
3: 13729 Honda
4: 2592660 Noriyuki_Haga
5: 15573 Japan
6: 332378 Yamaha_Motor_Company
7: 27816654 Wataru_Yoshikawa
8: 15573 Japan
9: 332378 Yamaha_Motor_Company

10: 2591706 Troy_Corser
11: 4689264 Australia
12: 8653 Ducati
13: 6249661 John_Kocinski
14: 3434750 United_States
15: 8653 Ducati
16: 6020107 Aaron_Slight
17: 4913064 New_Zealand
18: 13729 Honda
19: 15573 Japan
20: 332378 Yamaha_Motor_Company
21: 5307316 Carl_Fogarty
22: 31717 United_Kingdom
23: 13729 Honda
24: 7727384 Akira_Ryō
25: 15573 Japan
26: 59592 Kawasaki_Heavy_Industries
27: 15573 Japan
28: 59592 Kawasaki_Heavy_Industries
29: 2592660 Noriyuki_Haga
30: 15961118 Takuma_Aoki
31: 15573 Japan
32: 13729 Honda
33: 6249661 John_Kocinski
34: 2592660 Noriyuki_Haga
35: 6020107 Aaron_Slight
36: 5307316 Carl_Fogarty
37: 27816654 Wataru_Yoshikawa
38: 2591706 Troy_Corser
39: 15961118 Takuma_Aoki
40: 6020107 Aaron_Slight
41: 2591706 Troy_Corser
42: 6249661 John_Kocinski
43: 5307316 Carl_Fogarty
44: 2224069 Colin_Edwards
45: 3434750 United_States
46: 332378 Yamaha_Motor_Company
47: 14532 Italy
48: 8653 Ducati
49: 7459553 Simon_Crafar
50: 4913064 New_Zealand
51: 59592 Kawasaki_Heavy_Industries
52: 7452235 Anthony_Gobert
53: 4689264 Australia
54: 59592 Kawasaki_Heavy_Industries
55: 27816654 Wataru_Yoshikawa
56: 2885630 Neil_Hodgson
57: 31717 United_Kingdom
58: 8653 Ducati
59: 15573 Japan
60: 2592660 Noriyuki_Haga
61: 15961118 Takuma_Aoki

```
62: 15573 Japan
63: 13729 Honda
64: 6249661 John_Kocinski
65: 6020107 Aaron_Slight
66: 5307316 Carl_Fogarty
67: 27816654 Wataru_Yoshikawa
68: 2591706 Troy_Corser
69: 15573 Japan
70: 614614 Suzuki
71: 15961118 Takuma_Aoki
72: 6020107 Aaron_Slight
73: 2591706 Troy_Corser
74: 6249661 John_Kocinski
75: 5307316 Carl_Fogarty
76: 2224069 Colin_Edwards
77: 3434750 United_States
78: 332378 Yamaha_Motor_Company
79: 14532 Italy
80: 8653 Ducati
81: 7459553 Simon_Crafar
82: 4913064 New_Zealand
83: 59592 Kawasaki_Heavy_Industries
84: 7452235 Anthony_Gobert
85: 4689264 Australia
86: 59592 Kawasaki_Heavy_Industries
87: 27816654 Wataru_Yoshikawa
88: 2885630 Neil_Hodgson
89: 31717 United_Kingdom
90: 8653 Ducati
1: 15573 Japan
    [pem_pos=0, final_scores=0.6729421,
final_local_scores=0.75933605, pem=-0.13549899,
similarity=27.48668]
2: 15573 Japan
    [pem_pos=0, final_scores=0.6683418,
final_local_scores=0.7543109, pem=-0.13549899,
similarity=27.33672]
3: 13729 Honda
    [pem_pos=0, final_scores=0.46075785,
final_local_scores=0.4333036, pem=-0.17204832,
similarity=17.955982]
4: 2592660 Noriyuki_Haga
    [pem_pos=0, final_scores=0.7464022,
final_local_scores=0.7823456, pem=0.0, similarity=27.436466]
5: 15573 Japan
    [pem_pos=0, final_scores=0.73343927,
final_local_scores=0.82541996, pem=-0.13549899,
similarity=29.458755]
6: 332378 Yamaha_Motor_Company
    [pem_pos=1, final_scores=0.4839062,
final_local_scores=0.45767945, pem=-0.7066168,
similarity=21.590473]
```

7: 27816654 Wataru_Yoshikawa
[pem_pos=0, final_scores=0.6813215,
final_local_scores=0.7681463, pem=0.0, similarity=27.01273]
8: 15573 Japan
[pem_pos=0, final_scores=0.7378307,
final_local_scores=0.83021694, pem=-0.13549899,
similarity=29.601906]
9: 332378 Yamaha_Motor_Company
[pem_pos=1, final_scores=0.51966196,
final_local_scores=0.49673718, pem=-0.7066168,
similarity=22.756035]
10: 2591706 Troy_Corser
[pem_pos=0, final_scores=0.86121863,
final_local_scores=0.9010127, pem=0.0, similarity=30.97773]
11: 4689264 Australia
[pem_pos=0, final_scores=0.74374115,
final_local_scores=0.85165083, pem=-0.17168477,
similarity=30.43832]
12: 8653 Ducati
[pem_pos=0, final_scores=0.33144003,
final_local_scores=0.30483148, pem=0.0, similarity=13.186492]
13: 6249661 John_Kocinski
[pem_pos=0, final_scores=0.8176351,
final_local_scores=0.8618202, pem=0.0, similarity=29.808147]
14: 3434750 United_States
[pem_pos=0, final_scores=0.44085294,
final_local_scores=0.5197913, pem=-0.14075686,
similarity=20.366781]
15: 8653 Ducati
[pem_pos=0, final_scores=0.32563293,
final_local_scores=0.29848808, pem=0.0, similarity=12.997192]
16: 6020107 Aaron_Slight
[pem_pos=0, final_scores=0.6501724,
final_local_scores=0.69224167, pem=0.0, similarity=24.747585]
17: 4913064 New_Zealand
[pem_pos=0, final_scores=0.56287426,
final_local_scores=0.65268797, pem=-0.23795462,
similarity=24.861258]
18: 13729 Honda
[pem_pos=0, final_scores=0.531804,
final_local_scores=0.51091075, pem=-0.17204832,
similarity=20.271935]
19: 15573 Japan
[pem_pos=0, final_scores=0.76241225,
final_local_scores=0.8570686, pem=-0.13549899,
similarity=30.403214]
20: 332378 Yamaha_Motor_Company
[pem_pos=1, final_scores=0.5235931,
final_local_scores=0.50103134, pem=-0.7066168,
similarity=22.884178]
21: 5307316 Carl_Fogarty
[pem_pos=0, final_scores=0.75471884,

```
final_local_scores=0.77897394, pem=0.0, similarity=27.33585]
22: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.58172506,
final_local_scores=0.66683394, pem=-0.61114657,
similarity=27.312878]
23: 13729 Honda
    [pem_pos=0, final_scores=0.4780118,
final_local_scores=0.45215088, pem=-0.17204832,
similarity=18.518423]
24: 7727384 Akira_Ryō
    [pem_pos=0, final_scores=0.40842646,
final_local_scores=0.49648482, pem=0.0, similarity=18.90581]
25: 15573 Japan
    [pem_pos=0, final_scores=0.67994386,
final_local_scores=0.76698446, pem=-0.13549899,
similarity=27.714924]
26: 59592 Kawasaki_Heavy_Industries
    [pem_pos=0, final_scores=0.35706705,
final_local_scores=0.354334, pem=-0.58638686,
similarity=17.85261]
27: 15573 Japan
    [pem_pos=0, final_scores=0.7330387,
final_local_scores=0.8249825, pem=-0.13549899,
similarity=29.445702]
28: 59592 Kawasaki_Heavy_Industries
    [pem_pos=0, final_scores=0.3514459,
final_local_scores=0.34819376, pem=-0.58638686,
similarity=17.669373]
29: 2592660 Noriyuki_Haga
    [pem_pos=0, final_scores=1.2336885,
final_local_scores=1.3168375, pem=0.0, similarity=43.386772]
30: 15961118 Takuma_Aoki
    [pem_pos=0, final_scores=0.5676744,
final_local_scores=0.66213745, pem=0.0, similarity=23.849215]
31: 15573 Japan
    [pem_pos=0, final_scores=0.73825926,
final_local_scores=0.8306851, pem=-0.13549899,
similarity=29.615877]
32: 13729 Honda
    [pem_pos=0, final_scores=0.5043956,
final_local_scores=0.48097122, pem=-0.17204832,
similarity=19.378479]
33: 6249661 John_Kocinski
    [pem_pos=0, final_scores=0.7387746,
final_local_scores=0.7779157, pem=0.0, similarity=27.30427]
34: 2592660 Noriyuki_Haga
    [pem_pos=0, final_scores=1.0012633,
final_local_scores=1.0629482, pem=0.0, similarity=35.81021]
35: 6020107 Aaron_Slight
    [pem_pos=0, final_scores=0.8829177,
final_local_scores=0.94897693, pem=0.0, similarity=32.409077]
36: 5307316 Carl_Fogarty
```

[pem_pos=0, final_scores=1.0914465,
final_local_scores=1.1487503, pem=0.0, similarity=38.370716]
37: 27816654 Wataru_Yoshikawa
[pem_pos=0, final_scores=1.0963055,
final_local_scores=1.2246795, pem=0.0, similarity=40.63659]
38: 2591706 Troy_Corser
[pem_pos=0, final_scores=0.6770031,
final_local_scores=0.69978523, pem=0.0, similarity=24.9727]
39: 15961118 Takuma_Aoki
[pem_pos=0, final_scores=0.48433334,
final_local_scores=0.5711, pem=0.0, similarity=21.132477]
40: 6020107 Aaron_Slight
[pem_pos=0, final_scores=1.0492172,
final_local_scores=1.1306341, pem=0.0, similarity=37.830086]
41: 2591706 Troy_Corser
[pem_pos=0, final_scores=0.6960872,
final_local_scores=0.7206317, pem=0.0, similarity=25.5948]
42: 6249661 John_Kocinski
[pem_pos=0, final_scores=0.8456574,
final_local_scores=0.8946689, pem=0.0, similarity=30.788416]
43: 5307316 Carl_Fogarty
[pem_pos=0, final_scores=1.1149274,
final_local_scores=1.1743996, pem=0.0, similarity=39.136143]
44: 2224069 Colin_Edwards
[pem_pos=0, final_scores=0.5974936,
final_local_scores=0.6086402, pem=0.0, similarity=22.25275]
45: 3434750 United_States
[pem_pos=0, final_scores=0.46175992,
final_local_scores=0.54262906, pem=-0.14075686,
similarity=21.048304]
46: 332378 Yamaha_Motor_Company
[pem_pos=1, final_scores=0.47661203,
final_local_scores=0.44971162, pem=-0.7066168,
similarity=21.352697]
47: 14532 Italy
[pem_pos=0, final_scores=0.70920306,
final_local_scores=0.8137498, pem=-0.20612453,
similarity=29.494568]
48: 8653 Ducati
[pem_pos=0, final_scores=0.3802274,
final_local_scores=0.35812426, pem=0.0, similarity=14.776855]
49: 7459553 Simon_Crafar
[pem_pos=0, final_scores=0.5018618,
final_local_scores=0.5606487, pem=0.0, similarity=20.820587]
50: 4913064 New_Zealand
[pem_pos=0, final_scores=0.5685037,
final_local_scores=0.65883726, pem=-0.23795462,
similarity=25.044765]
51: 59592 Kawasaki_Heavy_Industries
[pem_pos=0, final_scores=0.35775417,
final_local_scores=0.3550846, pem=-0.58638686,
similarity=17.87501]

52: 7452235 Anthony_Gobert
[pem_pos=0, final_scores=0.66527736,
final_local_scores=0.71292883, pem=0.0, similarity=25.364931]
53: 4689264 Australia
[pem_pos=0, final_scores=0.6715556,
final_local_scores=0.7727991, pem=-0.17168477,
similarity=28.085228]
54: 59592 Kawasaki_Heavy_Industries
[pem_pos=0, final_scores=0.40898114,
final_local_scores=0.41104227, pem=-0.58638686,
similarity=19.544897]
55: 27816654 Wataru_Yoshikawa
[pem_pos=0, final_scores=1.1558166,
final_local_scores=1.2896862, pem=0.0, similarity=42.576523]
56: 2885630 Neil_Hodgson
[pem_pos=0, final_scores=0.605961,
final_local_scores=0.63151747, pem=0.0, similarity=22.935453]
57: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5230095,
final_local_scores=0.6026961, pem=-0.61114657,
similarity=25.398876]
58: 8653 Ducati
[pem_pos=0, final_scores=0.29136723,
final_local_scores=0.26105803, pem=0.0, similarity=11.880205]
59: 15573 Japan
[pem_pos=0, final_scores=0.76106185,
final_local_scores=0.8555935, pem=-0.99076414,
similarity=35.010258]
60: 2592660 Noriyuki_Haga
[pem_pos=0, final_scores=0.71408725,
final_local_scores=0.7470464, pem=0.0, similarity=26.383068]
61: 15961118 Takuma_Aoki
[pem_pos=0, final_scores=0.5263044,
final_local_scores=0.61694705, pem=0.0, similarity=22.500643]
62: 15573 Japan
[pem_pos=0, final_scores=0.740503,
final_local_scores=0.83313614, pem=-0.13549899,
similarity=29.68902]
63: 13729 Honda
[pem_pos=0, final_scores=0.4985745,
final_local_scores=0.47461253, pem=-0.17204832,
similarity=19.188723]
64: 6249661 John_Kocinski
[pem_pos=0, final_scores=0.73930544,
final_local_scores=0.77849555, pem=0.0, similarity=27.321573]
65: 6020107 Aaron_Slight
[pem_pos=0, final_scores=0.8520716,
final_local_scores=0.9152823, pem=0.0, similarity=31.403563]
66: 5307316 Carl_Fogarty
[pem_pos=0, final_scores=1.1105562,
final_local_scores=1.1696248, pem=0.0, similarity=38.993652]
67: 27816654 Wataru_Yoshikawa

[pem_pos=0, final_scores=1.1021873,
final_local_scores=1.2311044, pem=0.0, similarity=40.82832]
68: 2591706 Troy_Corser
[pem_pos=0, final_scores=0.72625893,
final_local_scores=0.7535897, pem=0.0, similarity=26.578333]
69: 15573 Japan
[pem_pos=0, final_scores=0.743938,
final_local_scores=0.8368883, pem=-0.13549899,
similarity=29.800993]
70: 614614 Suzuki
[pem_pos=0, final_scores=0.46256274,
final_local_scores=0.43819427, pem=-0.1481934,
similarity=17.972204]
71: 15961118 Takuma_Aoki
[pem_pos=0, final_scores=0.5036954,
final_local_scores=0.59225017, pem=0.0, similarity=21.76364]
72: 6020107 Aaron_Slight
[pem_pos=0, final_scores=1.051825,
final_local_scores=1.1334827, pem=0.0, similarity=37.9151]
73: 2591706 Troy_Corser
[pem_pos=0, final_scores=0.8122403,
final_local_scores=0.84751135, pem=0.0, similarity=29.381142]
74: 6249661 John_Kocinski
[pem_pos=0, final_scores=0.8510607,
final_local_scores=0.90057117, pem=0.0, similarity=30.964552]
75: 5307316 Carl_Fogarty
[pem_pos=0, final_scores=1.1073929,
final_local_scores=1.1661694, pem=0.0, similarity=38.890537]
76: 2224069 Colin_Edwards
[pem_pos=0, final_scores=0.6011588,
final_local_scores=0.61264384, pem=0.0, similarity=22.372227]
77: 3434750 United_States
[pem_pos=0, final_scores=0.46207446,
final_local_scores=0.5429726, pem=-0.14075686,
similarity=21.058556]
78: 332378 Yamaha_Motor_Company
[pem_pos=1, final_scores=0.47678292,
final_local_scores=0.4498983, pem=-0.7066168,
similarity=21.358269]
79: 14532 Italy
[pem_pos=0, final_scores=0.70924205,
final_local_scores=0.81379235, pem=-0.20612453,
similarity=29.495838]
80: 8653 Ducati
[pem_pos=0, final_scores=0.3795939,
final_local_scores=0.35743222, pem=0.0, similarity=14.756204]
81: 7459553 Simon_Crafar
[pem_pos=0, final_scores=0.49467432,
final_local_scores=0.55279744, pem=0.0, similarity=20.58629]
82: 4913064 New_Zealand
[pem_pos=0, final_scores=0.54447645,
final_local_scores=0.6325912, pem=-0.23795462,

```
similarity=24.26153]
83: 59592 Kawasaki_Heavy_Industries
    [pem_pos=0, final_scores=0.36484808,
final_local_scores=0.36283362, pem=-0.58638686,
similarity=18.106255]
84: 7452235 Anthony_Gobert
    [pem_pos=0, final_scores=0.6652787,
final_local_scores=0.7129303, pem=0.0, similarity=25.364975]
85: 4689264 Australia
    [pem_pos=0, final_scores=0.67528975,
final_local_scores=0.7768781, pem=-0.17168477,
similarity=28.206955]
86: 59592 Kawasaki_Heavy_Industries
    [pem_pos=0, final_scores=0.40786028,
final_local_scores=0.40981787, pem=-0.58638686,
similarity=19.50836]
87: 27816654 Wataru_Yoshikawa
    [pem_pos=0, final_scores=1.1560948,
final_local_scores=1.2899902, pem=0.0, similarity=42.58559]
88: 2885630 Neil_Hodgson
    [pem_pos=0, final_scores=0.5863521,
final_local_scores=0.6100977, pem=0.0, similarity=22.296246]
89: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.4942693,
final_local_scores=0.5713019, pem=-0.61114657,
similarity=24.46201]
90: 8653 Ducati
    [pem_pos=0, final_scores=0.2988931,
final_local_scores=0.26927888, pem=0.0, similarity=12.125532]
```

Sample: b'322_MOTORCYCLING'

MOTORCYCLING - [1 1 JAPANESE 1] WIN BOTH ROUND NINE
SUPERBIKE RACES . SUGO , [2 2 Japan 2] 1996-08-25 Teenager
Yuuichi Takeda , racing in only his first season , outclassed
the big names to win the first race in round nine of the
world superbike championship on Sunday . Takeda , 18 , showed
poise far beyond his years to overtake [3 3 Australian 3] [1 1
Ducati 1] rider [4 4 Troy Corser 4] , last year 's
championship runner-up , with four of the 25 laps left . [5 5
Honda 5] 's Takeda was pursued past [6 6 Corser 6] by the [7 7
Yamaha 7] duo of [8 8 Noriyuki Haga 8] and [9 9 Wataru Yoshik
awa 9] with [10 10 Haga 10] briefly taking the lead in the
final chicane on the last lap . But Takeda found one more
spurt of power to just take the flag first with [11 11 Haga 11]
] second and [12 12 Yoshikawa 12] third . [13 13 Haga 13] had
the consolation of recording the fastest lap at 147.159 kph .
[14 14 Corser 14] , who crashed during practice on Friday ,
limped in fourth , four seconds behind Takeda with
championship leader [15 15 Aaron Slight 15] of [16 16 New Zea
land 16] 11 seconds behind the winner . In the second race ,
Takeda again challenged strongly until the fifth lap from the
end when he crashed while running second to eventual race

winner [17 17 Takuma Aoki 17] . [18 18 Aoki 18] , the elder brother of reigning 125cc world champion [19 19 Haruchika 19] , had a race-long duel with [20 20 John Kocinski 20] of the [2 1 21 United States 21] on a [22 22 Ducati 22] before taking the chequered flag . [23 23 Kocinski 23] led for the early laps before he was passed first by [24 24 Aoki 24] , who recorded the fastest lap of 147.786 kph , and then by Takeda . With Takeda out of the race , [25 25 Kocinski 25] regained second place but he could not overtake [26 26 Aoki 26] . [27 27 Haga 27] again was the unlucky rider finishing third ahead of [28 28 Slight 28] with [29 29 Corser 29] in 10th place . But the strong showing by the [30 30 Japanese 30] riders did not alter the championship table with [31 31 Slight 31] still leading on 280 points , followed by [32 32 Corser 32] with 269 and [33 33 Kocinski 33] with 254 .

1: 15573 Japan
2: 15573 Japan
3: 4689264 Australia
4: 2591706 Troy_Corser
5: 6527836 Honda_Racing_Corporation
6: 2591706 Troy_Corser
7: 332378 Yamaha_Motor_Company
8: 2592660 Noriyuki_Haga
9: 27816654 Wataru_Yoshikawa
10: 2592660 Noriyuki_Haga
11: 2592660 Noriyuki_Haga
12: 27816654 Wataru_Yoshikawa
13: 2592660 Noriyuki_Haga
14: 2591706 Troy_Corser
15: 6020107 Aaron_Slight
16: 4913064 New_Zealand
17: 15961118 Takuma_Aoki
18: 15961118 Takuma_Aoki
19: 14836984 Haruchika_Aoki
20: 6249661 John_Kocinski
21: 3434750 United_States
22: 8653 Ducati
23: 6249661 John_Kocinski
24: 15961118 Takuma_Aoki
25: 6249661 John_Kocinski
26: 15961118 Takuma_Aoki
27: 2592660 Noriyuki_Haga
28: 6020107 Aaron_Slight
29: 2591706 Troy_Corser
30: 15573 Japan
31: 6020107 Aaron_Slight
32: 2591706 Troy_Corser
33: 6249661 John_Kocinski
1: 12080858 Ducati_Corse
[winner: 8653 Ducati, pem_pos=0, final_scores=0.11676256, final_local_scores=0.09048395, pem=0.0, similarity=6.7899346|
gt: 12080858 Ducati_Corse, pem_pos=1,

```
final_scores=0.06377815, final_local_scores=0.029067278,
pem=-1.4988651, similarity=13.1082 ]
1: 15573 Japan
    [pem_pos=0, final_scores=0.6596454,
final_local_scores=0.8017178, pem=-0.99076414,
similarity=33.4025]
2: 15573 Japan
    [pem_pos=0, final_scores=0.61766547,
final_local_scores=0.7466946, pem=-0.13549899,
similarity=27.109434]
3: 4689264 Australia
    [pem_pos=0, final_scores=0.5815388,
final_local_scores=0.72136277, pem=-0.491009,
similarity=28.2868]
4: 2591706 Troy_Corser
    [pem_pos=0, final_scores=0.72080266,
final_local_scores=0.7296569, pem=0.0, similarity=25.86413]
5: 6527836 Honda_Racing_Corporation
    [pem_pos=5, final_scores=0.34321493,
final_local_scores=0.35394692, pem=-3.9711916,
similarity=36.248154]
6: 2591706 Troy_Corser
    [pem_pos=0, final_scores=0.85484654,
final_local_scores=0.8760794, pem=0.0, similarity=30.233667]
7: 332378 Yamaha_Motor_Company
    [pem_pos=1, final_scores=0.39816433,
final_local_scores=0.37822193, pem=-0.7066168,
similarity=19.219303]
8: 2592660 Noriyuki_Haga
    [pem_pos=0, final_scores=0.8180061,
final_local_scores=0.8260321, pem=0.0, similarity=28.740158]
9: 27816654 Wataru_Yoshikawa
    [pem_pos=0, final_scores=0.6086183,
final_local_scores=0.69851273, pem=0.0, similarity=24.934727]
10: 2592660 Noriyuki_Haga
    [pem_pos=0, final_scores=1.1515017,
final_local_scores=1.1938374, pem=0.0, similarity=39.7162]
11: 2592660 Noriyuki_Haga
    [pem_pos=0, final_scores=1.2779075,
final_local_scores=1.3319166, pem=0.0, similarity=43.83676]
12: 27816654 Wataru_Yoshikawa
    [pem_pos=0, final_scores=0.97563726,
final_local_scores=1.107367, pem=0.0, similarity=37.135754]
13: 2592660 Noriyuki_Haga
    [pem_pos=0, final_scores=1.2333928,
final_local_scores=1.2832911, pem=0.0, similarity=42.38568]
14: 2591706 Troy_Corser
    [pem_pos=0, final_scores=0.81493133,
final_local_scores=0.8324781, pem=0.0, similarity=28.93252]
15: 6020107 Aaron_Slight
    [pem_pos=0, final_scores=0.7689902,
final_local_scores=0.80647093, pem=0.0, similarity=28.156414]
```

16: 4913064 New_Zealand
[pem_pos=0, final_scores=0.41693318,
final_local_scores=0.5376916, pem=-0.23795462,
similarity=21.429537]
17: 15961118 Takuma_Aoki
[pem_pos=0, final_scores=0.5485718,
final_local_scores=0.6287047, pem=0.0, similarity=22.851517]
18: 15961118 Takuma_Aoki
[pem_pos=0, final_scores=0.42245042,
final_local_scores=0.49093628, pem=0.0, similarity=18.74023]
19: 14836984 Haruchika_Aoki
[pem_pos=0, final_scores=0.37000906,
final_local_scores=0.42907107, pem=-0.6931472,
similarity=20.66349]
20: 6249661 John_Kocinski
[pem_pos=0, final_scores=0.82198995,
final_local_scores=0.84542954, pem=0.0, similarity=29.319017]
21: 3434750 United_States
[pem_pos=0, final_scores=0.4921866,
final_local_scores=0.62585866, pem=-0.073457584,
similarity=23.166058]
22: 8653 Ducati
[pem_pos=0, final_scores=0.2638604,
final_local_scores=0.24917844, pem=0.0, similarity=11.525695]
23: 6249661 John_Kocinski
[pem_pos=0, final_scores=0.8708895,
final_local_scores=0.90312773, pem=0.0, similarity=31.040846]
24: 15961118 Takuma_Aoki
[pem_pos=0, final_scores=0.5674335,
final_local_scores=0.64930826, pem=0.0, similarity=23.466366]
25: 6249661 John_Kocinski
[pem_pos=0, final_scores=0.9058917,
final_local_scores=0.9413623, pem=0.0, similarity=32.181843]
26: 15961118 Takuma_Aoki
[pem_pos=0, final_scores=0.581554,
final_local_scores=0.66473275, pem=0.0, similarity=23.926664]
27: 2592660 Noriyuki_Haga
[pem_pos=0, final_scores=1.2956895,
final_local_scores=1.3513408, pem=0.0, similarity=44.416416]
28: 6020107 Aaron_Slight
[pem_pos=0, final_scores=1.0433592,
final_local_scores=1.1113739, pem=0.0, similarity=37.25533]
29: 2591706 Troy_Corser
[pem_pos=0, final_scores=0.8738474,
final_local_scores=0.896835, pem=0.0, similarity=30.853058]
30: 15573 Japan
[pem_pos=0, final_scores=0.7844921,
final_local_scores=0.92892736, pem=-0.99076414,
similarity=37.19869]
31: 6020107 Aaron_Slight
[pem_pos=0, final_scores=1.0020369,
final_local_scores=1.0662355, pem=0.0, similarity=35.908306]

```
32: 2591706 Troy_Corser
    [pem_pos=0, final_scores=0.7223387,
final_local_scores=0.73133475, pem=0.0, similarity=25.9142]
33: 6249661 John_Kocinski
    [pem_pos=0, final_scores=0.90641636,
final_local_scores=0.9419355, pem=0.0, similarity=32.198948]
1: 8653 Ducati
    [pem_pos=0, final_scores=0.11676256,
final_local_scores=0.09048395, pem=0.0, similarity=6.7899346]
```

Sample: b'323_RUGBY'

```
[1 1 RUGBY LEAGUE 1] - EUROPEAN SUPER LEAGUE RESULTS /
STANDINGS . [1 1 LONDON 1] 1996-08-25 Results of [2 2 European
Super League 2] rugby league matches on Sunday : [3 3 Halifax
3] 64 [4 4 Leeds 4] 24 London 56 [5 5 Castleford 5] 0
Standings : [6 6 Wigan 6] 22 19 1 2 902 326 39 [7 7 St Helens
7] 21 19 0 2 884 441 38 [8 8 Bradford 8] 22 17 0 5 767 409 34
London 22 12 1 9 611 462 25 [9 9 Warrington 9] 21 12 0 9 555
499 24 [10 10 Halifax 10] 22 10 1 11 667 576 21 2 Sheffield
22 10 0 12 599 730 20 [11 11 Oldham 11] 22 9 1 12 473 681 19 [
12 12 Castleford 12] 22 9 0 13 548 599 18 [13 13 Leeds 13] 22
6 0 16 555 745 12 3 Paris 22 3 1 18 398 795 7 [14 14 Workingt
on 14] 22 2 1 19 325 1021 5
1: 17867 London
2: 22751937 Super_League
3: 1196374 Halifax_RLFC
4: 1095024 Leeds_Rhinos
5: 1216025 Castleford_Tigers
6: 1095029 Wigan_Warriors
7: 1095032 St_Helens_RLFC
8: 1095011 Bradford_Bulls
9: 852392 Warrington_Wolves
10: 1196374 Halifax_RLFC
11: 1098719 Oldham_Roughyeds
12: 1216025 Castleford_Tigers
13: 1095024 Leeds_Rhinos
14: 3646900 Workington_Town
1: 25402 Rugby_football
    [winner: 25735 Rugby_league, pem_pos=0,
final_scores=0.3185163, final_local_scores=0.29848585,
pem=0.0, similarity=12.997126]
    RECALL MISS: 25402 Rugby_football]
2: 3367769 Sheffield_Eagles
    [winner: 88885 Sheffield, pem_pos=0,
final_scores=-0.45116377, final_local_scores=-0.48158127,
pem=-0.25810194, similarity=-8.878035]
    RECALL MISS: 3367769 Sheffield_Eagles]
3: 5690457 Paris_Saint-Germain_Rugby_League
    [winner: 22989 Paris, pem_pos=0, final_scores=-0.5046057,
final_local_scores=-0.46655315, pem=-0.115301244,
similarity=-9.206138]
    RECALL MISS: 5690457 Paris_Saint-Germain_Rugby_League]
```

```
1: 17867 London
   [pem_pos=0, final_scores=0.54295754,
final_local_scores=0.6545907, pem=-0.09086737,
similarity=24.118156]
2: 22751937 Super_League
   [pem_pos=0, final_scores=0.47134268,
final_local_scores=0.4277084, pem=0.0, similarity=16.853384]
3: 1196374 Halifax_RLFC
   [pem_pos=3, final_scores=0.56852335,
final_local_scores=0.5142593, pem=-2.7129183,
similarity=34.189507]
4: 1095024 Leeds_Rhinos
   [pem_pos=1, final_scores=0.5023591,
final_local_scores=0.43720704, pem=-2.7286434,
similarity=31.97563]
5: 1216025 Castleford_Tigers
   [pem_pos=1, final_scores=0.5886317,
final_local_scores=0.5398511, pem=-0.78848183,
similarity=24.487835]
6: 1095029 Wigan_Warriors
   [pem_pos=1, final_scores=0.607535,
final_local_scores=0.56171685, pem=-1.8027086,
similarity=30.655872]
7: 1095032 St_Helens_RLFC
   [pem_pos=0, final_scores=0.7170759,
final_local_scores=0.67822343, pem=-1.1285291,
similarity=30.46637]
8: 1095011 Bradford_Bulls
   [pem_pos=9, final_scores=0.70474535,
final_local_scores=0.6612094, pem=-3.818427,
similarity=44.58672]
9: 852392 Warrington_Wolves
   [pem_pos=1, final_scores=0.7953382,
final_local_scores=0.76102704, pem=-1.6982093,
similarity=36.035404]
10: 1196374 Halifax_RLFC
    [pem_pos=3, final_scores=0.8482291,
final_local_scores=0.8197953, pem=-2.7129183,
similarity=43.307312]
11: 1098719 Oldham_Roughyeds
    [pem_pos=1, final_scores=0.58018965,
final_local_scores=0.53665227, pem=-2.0696447,
similarity=31.359535]
12: 1216025 Castleford_Tigers
    [pem_pos=1, final_scores=0.70663625,
final_local_scores=0.66875315, pem=-0.78848183,
similarity=28.334528]
13: 1095024 Leeds_Rhinos
    [pem_pos=1, final_scores=0.543809,
final_local_scores=0.48248476, pem=-2.7286434,
similarity=33.32681]
14: 3646900 Workington_Town
```

```
[pem_pos=3, final_scores=0.44156414,
final_local_scores=0.38974887, pem=-1.9682475,
similarity=26.424232]
1: 25735 Rugby_league
    [pem_pos=0, final_scores=0.3185163,
final_local_scores=0.29848585, pem=0.0, similarity=12.997126]
```

Sample: b'324_CRICKET'

CRICKET - [1 1 POLLOCK 1] CONCLUDES [2 2 WARWICKSHIRE 2] CAREER WITH FLOURISH . [3 3 LONDON 3] 1996-08-25 [4 4 South African 4] fast bowler [5 5 Shaun Pollock 5] concluded his [6 6 Warwickshire 6] career with a flourish on Sunday by taking the final three wickets during the county 's Sunday league victory over [7 7 Worcestershire 7] . [8 8 Pollock 8] , who returns home on Tuesday for an ankle operation , took the last three wickets in nine balls as [9 9 Worcestershire 9] were dismissed for 154 . After an hour 's interruption for rain , [10 10 Warwickshire 10] then reached an adjusted target of 109 with 13 balls to spare and record their fifth win in the last six games . [11 11 Warwickshire 11] are currently in fourth position behind [12 12 Yorkshire 12] , [13 13 Nottinghamshire 13] and [14 14 Surrey 14] . [15 15 Yorkshire 15] captain [16 16 David Byas 16] completed his third Sunday league century as his side swept clear at the top of the table , reaching a career best 111 not out against [17 17 Lancashire 17] . [18 18 Lancashire 18] 's total of 205 for eight from 40 overs looked reasonable before [19 19 Byas 19] put their attack to the sword , collecting his runs from just 100 balls with three sixes and nine fours . [20 20 Yorkshire 20] eventually reached their target with only four wickets down and 7.5 overs to spare .

```
1: 897244 Shaun_Pollock
2: 667292 Warwickshire_County_Cricket_Club
3: 17867 London
4: 17416221 South_Africa
5: 897244 Shaun_Pollock
6: 667292 Warwickshire_County_Cricket_Club
7: 1156428 Worcestershire_County_Cricket_Club
8: 897244 Shaun_Pollock
9: 1156428 Worcestershire_County_Cricket_Club
10: 667292 Warwickshire_County_Cricket_Club
11: 667292 Warwickshire_County_Cricket_Club
12: 869945 Yorkshire_County_Cricket_Club
13: 667341 Nottinghamshire_County_Cricket_Club
14: 766575 Surrey_County_Cricket_Club
15: 869945 Yorkshire_County_Cricket_Club
16: 2966347 David_Byas
17: 749919 Lancashire_County_Cricket_Club
18: 749919 Lancashire_County_Cricket_Club
19: 2966347 David_Byas
20: 869945 Yorkshire_County_Cricket_Club
1: 897244 Shaun_Pollock
```

```
[pem_pos=0, final_scores=1.3872694,
final_local_scores=1.4881985, pem=0.0, similarity=48.500523]
2: 667292 Warwickshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.74019974,
final_local_scores=0.7032504, pem=-1.1093602,
similarity=31.10898]
3: 17867 London
    [pem_pos=0, final_scores=0.57591087,
final_local_scores=0.67042613, pem=-0.09086737,
similarity=24.590715]
4: 17416221 South_Africa
    [pem_pos=0, final_scores=0.33115202,
final_local_scores=0.42719698, pem=-0.1389451,
similarity=17.593727]
5: 897244 Shaun_Pollock
    [pem_pos=0, final_scores=1.121201,
final_local_scores=1.1882513, pem=0.0, similarity=39.5495]
6: 667292 Warwickshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.61803836,
final_local_scores=0.56980765, pem=-1.1093602,
similarity=27.126783]
7: 1156428 Worcestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.66879994,
final_local_scores=0.63191116, pem=-0.9164197,
similarity=27.930836]
8: 897244 Shaun_Pollock
    [pem_pos=0, final_scores=1.3778317,
final_local_scores=1.4778893, pem=0.0, similarity=48.192875]
9: 1156428 Worcestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.7537624,
final_local_scores=0.7247198, pem=-0.9164197,
similarity=30.70043]
10: 667292 Warwickshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.73967725,
final_local_scores=0.70267963, pem=-1.1093602,
similarity=31.091948]
11: 667292 Warwickshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.78164476,
final_local_scores=0.74852276, pem=-1.1093602,
similarity=32.46]
12: 869945 Yorkshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6774269,
final_local_scores=0.6348005, pem=-1.7779728,
similarity=32.702316]
13: 667341 Nottinghamshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.5931466,
final_local_scores=0.54997414, pem=-1.1067253,
similarity=26.520582]
14: 766575 Surrey_County_Cricket_Club
    [pem_pos=1, final_scores=0.64744294,
final_local_scores=0.60892624, pem=-1.7952642,
similarity=32.02421]
```

```
15: 869945 Yorkshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.615648,
final_local_scores=0.5673165, pem=-1.7779728,
similarity=30.68846]
16: 2966347 David_Byas
    [pem_pos=0, final_scores=0.50720626,
final_local_scores=0.571814, pem=0.0, similarity=21.153784]
17: 749919 Lancashire_County_Cricket_Club
    [pem_pos=1, final_scores=0.5526182,
final_local_scores=0.50758415, pem=-1.7295884,
similarity=28.642807]
18: 749919 Lancashire_County_Cricket_Club
    [pem_pos=1, final_scores=0.63206726,
final_local_scores=0.5943701, pem=-1.7295884,
similarity=31.232676]
19: 2966347 David_Byas
    [pem_pos=0, final_scores=0.7055201,
final_local_scores=0.8018715, pem=0.0, similarity=28.019157]
20: 869945 Yorkshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.7505904,
final_local_scores=0.71472055, pem=-1.7779728,
similarity=35.08729]
```

Sample: b'325_SOCCER'

SOCCER - [1 1 HINCHCLIFFE 1] CALLED INTO [2 2 ENGLAND 2]
SQUAD . [3 3 LONDON 3] 1996-08-25 [4 4 England 4] manager [5 5
Glenn Hoddle 5] called up uncapped [6 6 Everton 6] defender [7
7 Andy Hinchcliffe 7] on Sunday to the national squad for the
opening [8 8 World Cup 8] qualifier against [9 9 Moldova 9]
next weekend . Left-back [10 10 Hinchcliffe 10] , 27 ,
replaces [11 11 Tottenham 11] 's [12 12 Darren Anderton 12]
who has a recurring groin problem .

```
1: 3733975 Andy_Hinchcliffe
2: 9316 England
3: 17867 London
4: 9316 England
5: 552843 Glenn_Hoddle
6: 91155 Everton_F.C.
7: 3733975 Andy_Hinchcliffe
8: 126460 1998_FIFA_World_Cup
9: 1011292 Moldova_national_football_team
10: 3733975 Andy_Hinchcliffe
11: 68198 Tottenham_Hotspur_F.C.
12: 1089164 Darren_Anderton
1: 3733975 Andy_Hinchcliffe
    [pem_pos=0, final_scores=0.60077715,
final_local_scores=0.6048576, pem=0.0, similarity=22.139872]
2: 9316 England
    [pem_pos=0, final_scores=0.44333524,
final_local_scores=0.49409264, pem=-0.19053856,
similarity=19.870602]
3: 17867 London
```

```
[pem_pos=0, final_scores=0.6516736,
final_local_scores=0.7215589, pem=-0.09086737,
similarity=26.116621]
4: 9316 England
    [pem_pos=0, final_scores=0.2941473,
final_local_scores=0.33112758, pem=-0.19053856,
similarity=15.007399]
5: 552843 Glenn_Hoddle
    [pem_pos=0, final_scores=0.5797104,
final_local_scores=0.59176195, pem=0.0, similarity=21.74907]
6: 91155 Everton_F.C.
    [pem_pos=0, final_scores=0.5655105,
final_local_scores=0.53480446, pem=-0.17171584,
similarity=20.983164]
7: 3733975 Andy_Hinchcliffe
    [pem_pos=0, final_scores=0.66339785,
final_local_scores=0.6619202, pem=0.0, similarity=23.842731]
8: 126460 1998_FIFA_World_Cup
    [pem_pos=18, final_scores=0.32071644,
final_local_scores=0.3315804, pem=-4.2172847,
similarity=36.918983]
9: 1011292 Moldova_national_football_team
    [pem_pos=1, final_scores=0.45621413,
final_local_scores=0.5053829, pem=-2.4170835,
similarity=32.315823]
10: 3733975 Andy_Hinchcliffe
    [pem_pos=0, final_scores=0.76867837,
final_local_scores=0.788677, pem=0.0, similarity=27.625406]
11: 68198 Tottenham_Hotspur_F.C.
    [pem_pos=1, final_scores=0.39088434,
final_local_scores=0.34727687, pem=-0.99701256,
similarity=19.875057]
12: 1089164 Darren_Anderton
    [pem_pos=0, final_scores=0.6176547,
final_local_scores=0.66367507, pem=0.0, similarity=23.895102]
```

Sample: b'326_CRICKET'

```
CRICKET - [1 1 ENGLAND 1] 326 AND 74-0 ; [2 2 PAKISTAN 2]
521-8 DECLARED . [3 3 LONDON 3] 1996-08-25 [4 4 England 4]
were 74 for no wicket in their second innings at the close of
the fourth day of the third and final test at [5 5 The Oval 5]
on Sunday . Scores : [6 6 England 6] 326 and 74-0 ; [7 7 Paki
stan 7] 521-8 declared .
1: 759125 England_cricket_team
2: 759295 Pakistan_national_cricket_team
3: 17867 London
4: 759125 England_cricket_team
5: 387687 The_Oval
6: 759125 England_cricket_team
7: 759295 Pakistan_national_cricket_team
1: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.6948356,
```

```
final_local_scores=0.6453319, pem=-3.3054674,
similarity=41.32335]
2: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7495374,
final_local_scores=0.72882766, pem=-2.7556365,
similarity=40.824963]
3: 17867 London
    [pem_pos=0, final_scores=0.55388314,
final_local_scores=0.6326511, pem=-0.09086737,
similarity=23.463434]
4: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.7224904,
final_local_scores=0.6755405, pem=-3.3054674,
similarity=42.224834]
5: 387687 The_Oval
    [pem_pos=0, final_scores=0.66345793,
final_local_scores=0.65478843, pem=0.0, similarity=23.629906]
6: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.7485263,
final_local_scores=0.70398074, pem=-3.3054674,
similarity=43.073547]
7: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.81055444,
final_local_scores=0.7954795, pem=-2.7556365,
similarity=42.813988]
```

Sample: b'327_SOCCER'
SOCCER - [1 1 ENGLISH 1] PREMIER LEAGUE SUMMARY . [2 2 LONDON
2] 1996-08-25 Summary of an [3 3 English 3] premier league
soccer match on Sunday : [4 4 Manchester United 4] 2 ([5 5 Cr
uyff 5] 39th minute , [6 6 Solskjaer 6] 70th) [7 7 Blackburn
7] 2 ([8 8 Warhurst 8] 34th , [9 9 Bohinen 9] 51st) .
Halftime 1-1 . Attendance 54,178 .

```
1: 9316 England
2: 17867 London
3: 9316 England
4: 19961 Manchester_United_F.C.
5: 1771982 Jordi_Cruyff
6: 194455 Ole_Gunnar_Solskjær
7: 298599 Blackburn_Rovers_F.C.
8: 3385586 Paul_Warhurst
9: 1882023 Lars_Bohinen
1: 9316 England
    [pem_pos=1, final_scores=0.50265855,
final_local_scores=0.54476506, pem=-1.3527626,
similarity=27.703121]
2: 17867 London
    [pem_pos=0, final_scores=0.66446155,
final_local_scores=0.72516716, pem=-0.09086737,
similarity=26.224297]
3: 9316 England
    [pem_pos=1, final_scores=0.6237518,
```

```

final_local_scores=0.6770411, pem=-1.3527626,
similarity=31.650501]
4: 19961 Manchester_United_F.C.
    [pem_pos=0, final_scores=0.44295412,
final_local_scores=0.3938285, pem=0.0, similarity=15.842341]
5: 1771982 Jordi_Cruyff
    [pem_pos=1, final_scores=0.2716145,
final_local_scores=0.29895243, pem=-0.45332113,
similarity=15.47628]
6: 194455 Ole_Gunnar_Solskjær
    [pem_pos=0, final_scores=0.31875247,
final_local_scores=0.31875324, pem=0.0, similarity=13.601946]
7: 298599 Blackburn_Rovers_F.C.
    [pem_pos=1, final_scores=0.7270153,
final_local_scores=0.7083836, pem=-2.734066,
similarity=40.09757]
8: 3385586 Paul_Warhurst
    [pem_pos=0, final_scores=0.45586443,
final_local_scores=0.4912852, pem=-0.025975484,
similarity=18.891903]
9: 1882023 Lars_Bohinne
    [pem_pos=0, final_scores=0.3978616,
final_local_scores=0.45402163, pem=0.0, similarity=17.638624]

```

Sample: b'328_SOCCER'

```

SOCCER - [1 1 ENGLISH 1] LEAGUE RESULTS / STANDINGS . [2 2 L0
NDON 2] 1996-08-25 Results of [3 3 English 3] league soccer
matches on Sunday : Premier league [4 4 Manchester United 4]
2 [5 5 Blackburn 5] 2 Standings ( ( tabulated under played ,
won , drawn , lost , goals for , against , points ) : [6 6 Sh
effield Wednesday 6] 3 3 0 0 6 2 9 [7 7 Chelsea 7] 3 2 1 0 3
0 7 [8 8 Arsenal 8] 3 2 0 1 4 2 6 [9 9 Aston Villa 9] 3 2 0 1
4 2 6 [10 10 Manchester United 10] 3 1 2 0 7 4 5 [11 11 Sunde
rland 11] 3 1 2 0 4 1 5 [12 12 Liverpool 12] 3 1 2 0 5 3 5 [13
13 Everton 13] 3 1 2 0 4 2 5 [14 14 Tottenham 14] 3 1 2 0 3 1
5 [15 15 Nottingham Forest 15] 3 1 1 1 5 5 4 [16 16 West Ham 1
6] 3 1 1 1 3 4 4 [17 17 Leicester 17] 3 1 1 1 2 3 4 [18 18 Ne
wcastle 18] 3 1 0 2 3 4 3 [19 19 Middlesbrough 19] 3 0 2 1 4
5 2 [20 20 Derby 20] 3 0 2 1 4 6 2 [21 21 Leeds 21] 2 0 1 1 3
5 1 [22 22 Southampton 22] 3 0 1 2 2 4 1 [23 23 Blackburn 23]
3 0 1 2 2 5 1 [24 24 Coventry 24] 3 0 1 2 1 6 1 [25 25 Wimble
don 25] 2 0 0 2 0 5 0 Division one [26 26 Barnsley 26] 3 [27 2
7 Huddersfield 27] 1 Standings : [28 28 Bolton 28] 3 2 1 0 5
2 7 [29 29 Barnsley 29] 2 2 0 0 5 2 6 [30 30 Wolverhampton 30]
2 2 0 0 4 1 6 [31 31 Queens Park Rangers 31] 2 2 0 0 4 2 6 [32
32 Stoke 32] 2 2 0 0 4 2 6 [33 33 Birmingham 33] 2 1 1 0 5 4
4 [34 34 Tranmere 34] 2 1 1 0 4 3 4 [35 35 Oxford 35] 2 1 0 1
6 2 3 [36 36 Ipswich 36] 2 1 0 1 5 3 3 1 Bradford 2 1 0 1 3 2
3 [37 37 Crystal Palace 37] 2 1 0 1 3 2 3 [38 38 Huddersfield
38] 2 1 0 1 3 3 3 [39 39 Norwich 39] 2 1 0 1 3 3 3 [40 40 Rea
ding 40] 2 1 0 1 3 5 3 [41 41 Manchester City 41] 3 1 0 2 2 3
3 [42 42 Port Vale 42] 2 0 2 0 2 2 2 [43 43 Sheffield United 4

```

[3] 2 0 1 1 4 5 1 [44 44 West Bromwich 44] 2 0 1 1 2 3 1 [45 45
Charlton 45] 2 0 1 1 1 3 1 [46 46 Swindon 46] 2 0 1 1 1 3 1 [4
7 47 Southend 47] 2 0 1 1 1 6 1 [48 48 Grimsby 48] 2 0 0 2 3
6 0 [49 49 Oldham 49] 2 0 0 2 2 5 0 [50 50 Portsmouth 50] 2 0
0 2 2 5 0

1: 9316 England
2: 17867 London
3: 9316 England
4: 19961 Manchester_United_F.C.
5: 298599 Blackburn_Rovers_F.C.
6: 177538 Sheffield_Wednesday_F.C.
7: 7473 Chelsea_F.C.
8: 2174 Arsenal_F.C.
9: 42173 Aston_Villa_F.C.
10: 19961 Manchester_United_F.C.
11: 184474 Sunderland_A.F.C.
12: 18119 Liverpool_F.C.
13: 91155 Everton_F.C.
14: 68198 Tottenham_Hotspur_F.C.
15: 157294 Nottingham_Forest_F.C.
16: 46417 West_Ham_United_F.C.
17: 298602 Leicester_City_F.C.
18: 57802 Newcastle_United_F.C.
19: 298603 Middlesbrough_F.C.
20: 4721303 Derby_County_F.C.
21: 7609513 Leeds_United_A.F.C.
22: 81576 Southampton_F.C.
23: 298599 Blackburn_Rovers_F.C.
24: 435021 Coventry_City_F.C.
25: 740757 Wimbledon_F.C.
26: 333188 Barnsley_F.C.
27: 236211 Huddersfield_Town_F.C.
28: 298601 Bolton_Wanderers_F.C.
29: 333188 Barnsley_F.C.
30: 217730 Wolverhampton_Wanderers_F.C.
31: 84588 Queens_Park_Rangers_F.C.
32: 203434 Stoke_City_F.C.
33: 185196 Birmingham_City_F.C.
34: 69219 Tranmere_Rovers_F.C.
35: 451165 Oxford_United_F.C.
36: 352843 Ipswich_Town_F.C.
37: 385313 Crystal_Palace_F.C.
38: 236211 Huddersfield_Town_F.C.
39: 248083 Norwich_City_F.C.
40: 168007 Reading_F.C.
41: 165813 Manchester_City_F.C.
42: 203456 Port_Vale_F.C.
43: 314933 Sheffield_United_F.C.
44: 33921 West_Bromwich_Albion_F.C.
45: 6720 Charlton_Athletic_F.C.
46: 74291 Swindon_Town_F.C.
47: 451168 Southend_United_F.C.

48: 451129 Grimsby_Town_F.C.
49: 451131 Oldham_Athletic_A.F.C.
50: 564301 Portsmouth_F.C.
1: 431144 Bradford_City_A.F.C.
[winner: 1095011 Bradford_Bulls, pem_pos=9,
final_scores=-0.34596932, final_local_scores=-0.4091799,
pem=-3.818427, similarity=12.644169]
RECALL MISS: 431144 Bradford_City_A.F.C.]
1: 9316 England
[pem_pos=1, final_scores=0.4911275,
final_local_scores=0.5513472, pem=-1.3527626,
similarity=27.899546]
2: 17867 London
[pem_pos=0, final_scores=0.6280818,
final_local_scores=0.7109813, pem=-0.09086737,
similarity=25.800964]
3: 9316 England
[pem_pos=1, final_scores=0.6353575,
final_local_scores=0.7088966, pem=-1.3527626,
similarity=32.60113]
4: 19961 Manchester_United_F.C.
[pem_pos=0, final_scores=0.43115407,
final_local_scores=0.36625117, pem=0.0, similarity=15.019378]
5: 298599 Blackburn_Rovers_F.C.
[pem_pos=1, final_scores=0.5560008,
final_local_scores=0.48915166, pem=-2.734066,
similarity=33.55525]
6: 177538 Sheffield_Wednesday_F.C.
[pem_pos=0, final_scores=0.3310029,
final_local_scores=0.2689927, pem=0.0, similarity=12.116991]
7: 7473 Chelsea_F.C.
[pem_pos=0, final_scores=0.5865174,
final_local_scores=0.54145956, pem=-0.51707125,
similarity=23.05986]
8: 2174 Arsenal_F.C.
[pem_pos=0, final_scores=0.5407942,
final_local_scores=0.4825856, pem=-0.15820743,
similarity=19.351387]
9: 42173 Aston_Villa_F.C.
[pem_pos=0, final_scores=0.48478717,
final_local_scores=0.4134493, pem=0.0, similarity=16.427864]
10: 19961 Manchester_United_F.C.
[pem_pos=0, final_scores=0.418307,
final_local_scores=0.35221767, pem=0.0, similarity=14.600591]
11: 184474 Sunderland_A.F.C.
[pem_pos=0, final_scores=0.727001,
final_local_scores=0.678526, pem=-0.39640337,
similarity=26.493986]
12: 18119 Liverpool_F.C.
[pem_pos=1, final_scores=0.6689781,
final_local_scores=0.6237598, pem=-1.1702383,
similarity=29.067888]

13: 91155 Everton_F.C.
[pem_pos=0, final_scores=0.51489705,
final_local_scores=0.44783765, pem=-0.17171584,
similarity=18.3879]
14: 68198 Tottenham_Hotspur_F.C.
[pem_pos=1, final_scores=0.41604567,
final_local_scores=0.34256417, pem=-0.99701256,
similarity=19.73442]
15: 157294 Nottingham_Forest_F.C.
[pem_pos=0, final_scores=0.5066578,
final_local_scores=0.43728828, pem=0.0, similarity=17.139267]
16: 46417 West_Ham_United_F.C.
[pem_pos=0, final_scores=0.689377,
final_local_scores=0.63913983, pem=-0.5127755,
similarity=25.951473]
17: 298602 Leicester_City_F.C.
[pem_pos=3, final_scores=0.943633,
final_local_scores=0.9153686, pem=-3.2133117,
similarity=48.880627]
18: 57802 Newcastle_United_F.C.
[pem_pos=2, final_scores=0.6906448,
final_local_scores=0.6382415, pem=-2.4885974,
similarity=36.66949]
19: 298603 Middlesbrough_F.C.
[pem_pos=0, final_scores=0.6159777,
final_local_scores=0.55798274, pem=-0.5175458,
similarity=23.555525]
20: 4721303 Derby_County_F.C.
[pem_pos=5, final_scores=0.81293255,
final_local_scores=0.7725443, pem=-3.703934,
similarity=47.286545]
21: 7609513 Leeds_United_A.F.C.
[pem_pos=4, final_scores=0.7576956,
final_local_scores=0.7117999, pem=-3.421221,
similarity=43.936375]
22: 81576 Southampton_F.C.
[pem_pos=1, final_scores=0.8251392,
final_local_scores=0.79631084, pem=-1.1503892,
similarity=34.10921]
23: 298599 Blackburn_Rovers_F.C.
[pem_pos=1, final_scores=0.68639296,
final_local_scores=0.6315853, pem=-2.734066,
similarity=37.805756]
24: 435021 Coventry_City_F.C.
[pem_pos=2, final_scores=0.69647294,
final_local_scores=0.6448614, pem=-3.1334689,
similarity=40.373955]
25: 740757 Wimbledon_F.C.
[pem_pos=1, final_scores=0.6887399,
final_local_scores=0.6561387, pem=-1.8899087,
similarity=33.94782]
26: 333188 Barnsley_F.C.

```
[pem_pos=0, final_scores=0.4244185,
final_local_scores=0.34998527, pem=-0.7126868,
similarity=18.409674]
27: 236211 Huddersfield_Town_F.C.
    [pem_pos=3, final_scores=0.8534525,
final_local_scores=0.81827515, pem=-2.2859883,
similarity=40.940235]
28: 298601 Bolton_Wanderers_F.C.
    [pem_pos=2, final_scores=0.571207,
final_local_scores=0.506314, pem=-2.2845325,
similarity=31.622776]
29: 333188 Barnsley_F.C.
    [pem_pos=0, final_scores=0.39121336,
final_local_scores=0.31371364, pem=-0.7126868,
similarity=17.327255]
30: 217730 Wolverhampton_Wanderers_F.C.
    [pem_pos=1, final_scores=0.8168033,
final_local_scores=0.773955, pem=-2.274954,
similarity=39.55763]
31: 84588 Queens_Park_Rangers_F.C.
    [pem_pos=0, final_scores=0.49273747,
final_local_scores=0.42434907, pem=0.0, similarity=16.753136]
32: 203434 Stoke_City_F.C.
    [pem_pos=0, final_scores=0.53343946,
final_local_scores=0.47227824, pem=-0.3417574,
similarity=20.041967]
33: 185196 Birmingham_City_F.C.
    [pem_pos=2, final_scores=0.7255253,
final_local_scores=0.6772042, pem=-3.102209,
similarity=41.169132]
34: 69219 Tranmere_Rovers_F.C.
    [pem_pos=0, final_scores=0.44712156,
final_local_scores=0.38526767, pem=-0.6002322,
similarity=18.851025]
35: 451165 Oxford_United_F.C.
    [pem_pos=18, final_scores=0.9814709,
final_local_scores=0.96784693, pem=-4.3872933,
similarity=56.830982]
36: 352843 Ipswich_Town_F.C.
    [pem_pos=4, final_scores=0.8149325,
final_local_scores=0.7792977, pem=-2.6931517,
similarity=41.991287]
37: 385313 Crystal_Palace_F.C.
    [pem_pos=0, final_scores=0.5186501,
final_local_scores=0.4514721, pem=-0.51090026,
similarity=20.340895]
38: 236211 Huddersfield_Town_F.C.
    [pem_pos=3, final_scores=0.72299135,
final_local_scores=0.67576617, pem=-2.2859883,
similarity=36.68748]
39: 248083 Norwich_City_F.C.
    [pem_pos=2, final_scores=0.8885279,
```

```
final_local_scores=0.85498124, pem=-2.7278075,
similarity=44.4383]
40: 168007 Reading_F.C.
    [pem_pos=0, final_scores=1.020937,
final_local_scores=1.0013354, pem=-1.2007197,
similarity=40.50126]
41: 165813 Manchester_City_F.C.
    [pem_pos=0, final_scores=0.34911352,
final_local_scores=0.27111685, pem=0.0, similarity=12.180381]
42: 203456 Port_Vale_F.C.
    [pem_pos=0, final_scores=0.57295203,
final_local_scores=0.5269976, pem=0.0, similarity=19.816374]
43: 314933 Sheffield_United_F.C.
    [pem_pos=0, final_scores=0.34729433,
final_local_scores=0.26376668, pem=0.0, similarity=11.961037]
44: 33921 West_Bromwich_Albion_F.C.
    [pem_pos=2, final_scores=0.593504,
final_local_scores=0.533577, pem=-1.5002033,
similarity=28.171055]
45: 6720 Charlton_Athletic_F.C.
    [pem_pos=1, final_scores=0.78795797,
final_local_scores=0.7500281, pem=-2.103781,
similarity=37.912735]
46: 74291 Swindon_Town_F.C.
    [pem_pos=4, final_scores=0.81816393,
final_local_scores=0.7846665, pem=-2.489671,
similarity=41.044945]
47: 451168 Southend_United_F.C.
    [pem_pos=2, final_scores=0.7369221,
final_local_scores=0.69636506, pem=-1.771178,
similarity=34.502575]
48: 451129 Grimsby_Town_F.C.
    [pem_pos=2, final_scores=0.828849,
final_local_scores=0.7962778, pem=-2.6069987,
similarity=42.029495]
49: 451131 Oldham_Athletic_A.F.C.
    [pem_pos=3, final_scores=0.6801136,
final_local_scores=0.6313012, pem=-2.8632221,
similarity=38.49965]
50: 564301 Portsmouth_F.C.
    [pem_pos=1, final_scores=0.7390062,
final_local_scores=0.6924153, pem=-1.2302613,
similarity=31.44312]
```

Sample: b'329_CRICKET'

CRICKET - [1 1 ENGLAND 1] V [2 2 PAKISTAN 2] FINAL TEST
SCOREBOARD . [3 3 LONDON 3] 1996-08-25 Scoreboard on the
fourth day of the third and final test between [4 4 England 4]
and [5 5 Pakistan 5] at The [6 6 Oval 6] on Sunday : [7 7 Eng
land 7] first innings 326 (J. Crawley 106 , G. Thorpe 54 ; [8
8 Waqar Younis 8] 4-95) [9 9 Pakistan 9] first innings
(overnight 339-4) [10 10 Saeed Anwar 10] c [11 11 Croft 11]

b [12 12 Cork 12] 176 Aamir Sohail c [13 13 Cork 13] b [14 14 Croft 14] 46 [15 15 Ijaz Ahmed 15] c 1 Stewart b [16 16 Mulla lly 16] 61 [17 17 Inzamam-ul-Haq 17] c [18 18 Hussain 18] b [1 9 19 Mullally 19] 35 [20 20 Salim Malik 20] not out 100 [21 21 Asif Mujtaba 21] run out 13 [22 22 Wasim Akram 22] st 2 Stewa rt b [23 23 Croft 23] 40 [24 24 Moin Khan 24] b 3 Salisbury 23 [25 25 Mushtaq Ahmed 25] c 4 Crawley b [26 26 Mullally 26] 2 [27 27 Waqar Younis 27] not out 0 Extras (b-4 lb-5 nb-16) 25 Total (for eight wickets , declared) 521 Fall of wickets : 1-106 2-239 3-334 4-334 5-365 6-440 7-502 8-519 Did not bat : Mohammad Akam Bowling : Lewis 23-3-112-0 , [28 28 Mu llally 28] 37.1-7-97-3 , [29 29 Croft 29] 47-10-116-2 , [30 30 Cork 30] 23-5-71-1 , 5 Salisbury 29-3-116-1 [31 31 England 31] second innings M. Atherton not out 26 A. Stewart not out 40 Extras (nb-8) 8 Total (for no wicket) 74 Bowling (to date) : [32 32 Wasim Akram 32] 7-0-35-0 , [33 33 Waqar Younis 33] 7-1-24-0 , [34 34 Mushtaq Ahmed 34] 7-2-11-0 , Aamir Sohail 2-1-4-0

1: 9316 England
2: 23235 Pakistan
3: 17867 London
4: 759125 England_cricket_team
5: 759295 Pakistan_national_cricket_team
6: 387687 The_Oval
7: 759125 England_cricket_team
8: 894433 Waqar_Younis
9: 759295 Pakistan_national_cricket_team
10: 894437 Saeed_Anwar
11: 1160680 Robert_Croft
12: 2036840 Dominic_Cork
13: 2036840 Dominic_Cork
14: 1160680 Robert_Croft
15: 938863 Ijaz_Ahmed_(cricketer)
16: 2222543 Alan_Mullally
17: 896403 Inzamam-ul-Haq
18: 682925 Nasser_Hussain
19: 2222543 Alan_Mullally
20: 938861 Saleem_Malik
21: 3976606 Asif_Mujtaba
22: 501393 Wasim_Akram
23: 1160680 Robert_Croft
24: 554132 Moin_Khan
25: 896405 Mushtaq_Ahmed
26: 2222543 Alan_Mullally
27: 894433 Waqar_Younis
28: 2222543 Alan_Mullally
29: 1160680 Robert_Croft
30: 2036840 Dominic_Cork
31: 759125 England_cricket_team
32: 501393 Wasim_Akram
33: 894433 Waqar_Younis
34: 896405 Mushtaq_Ahmed

1: 896414 Alec_Stewart
[winner: 5608588 Alexander_David_Stewart, pem_pos=0,
final_scores=-0.72572726, final_local_scores=-0.57711244,
pem=0.0, similarity=-13.132472]
RECALL MISS: 896414 Alec_Stewart]

2: 896414 Alec_Stewart
[winner: 292988 Al_Stewart, pem_pos=1,
final_scores=-0.6094884, final_local_scores=-0.53697383,
pem=0.0, similarity=-11.934656]
RECALL MISS: 896414 Alec_Stewart]

3: 2416655 Ian_Salisbury
[winner: 72119 Salisbury, pem_pos=0,
final_scores=-0.4940396, final_local_scores=-0.48847055,
pem=-0.80015767, similarity=-6.1358414]
RECALL MISS: 2416655 Ian_Salisbury]

4: 2930949 John_Crawley
[winner: 156618 Crawley, pem_pos=0,
final_scores=-0.49174583, final_local_scores=-0.48131818,
pem=-0.047762234, similarity=-10.014043]
RECALL MISS: 2930949 John_Crawley]

5: 2416655 Ian_Salisbury
[winner: 72119 Salisbury, pem_pos=0,
final_scores=-0.513419, final_local_scores=-0.5096396,
pem=-0.80015767, similarity=-6.7675686]
RECALL MISS: 2416655 Ian_Salisbury]

1: 9316 England
[pem_pos=0, final_scores=0.3803479,
final_local_scores=0.4648068, pem=-0.19053856,
similarity=18.996653]

2: 23235 Pakistan
[pem_pos=0, final_scores=0.29875332,
final_local_scores=0.38031358, pem=-0.21944217,
similarity=16.63239]

3: 17867 London
[pem_pos=0, final_scores=0.6044315,
final_local_scores=0.72267693, pem=-0.09086737,
similarity=26.149984]

4: 759125 England_cricket_team
[pem_pos=2, final_scores=0.8115433,
final_local_scores=0.7938861, pem=-3.3054674,
similarity=45.756504]

5: 759295 Pakistan_national_cricket_team
[pem_pos=1, final_scores=0.8375643,
final_local_scores=0.8272515, pem=-2.7556365,
similarity=43.762127]

6: 387687 The_Oval
[pem_pos=3, final_scores=0.39776367,
final_local_scores=0.37222737, pem=-2.1624494,
similarity=26.957457]

7: 759125 England_cricket_team
[pem_pos=2, final_scores=0.8275673,
final_local_scores=0.81139, pem=-3.3054674,

```
similarity=46.278854]
8: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.69982463,
final_local_scores=0.6945051, pem=0.0, similarity=24.81513]
9: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.86664575,
final_local_scores=0.8590186, pem=-2.7556365,
similarity=44.71012]
10: 894437 Saeed_Anwar
    [pem_pos=0, final_scores=0.79386824,
final_local_scores=0.8073589, pem=0.0, similarity=28.182913]
11: 1160680 Robert_Croft
    [pem_pos=7, final_scores=1.0067194,
final_local_scores=1.0471101, pem=-3.16984,
similarity=52.575653]
12: 2036840 Dominic_Cork
    [pem_pos=8, final_scores=0.9140983,
final_local_scores=0.926084, pem=-3.606134,
similarity=51.336624]
13: 2036840 Dominic_Cork
    [pem_pos=8, final_scores=1.1430135,
final_local_scores=1.1761391, pem=-3.606134,
similarity=58.79877]
14: 1160680 Robert_Croft
    [pem_pos=7, final_scores=0.99979,
final_local_scores=1.0395408, pem=-3.16984,
similarity=52.34977]
15: 938863 Ijaz_Ahmed_(cricketer)
    [pem_pos=0, final_scores=0.61831564,
final_local_scores=0.6212176, pem=0.0, similarity=22.628086]
16: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.77103525,
final_local_scores=0.79017174, pem=-0.73930055,
similarity=31.690447]
17: 896403 Inzamam-ul-Haq
    [pem_pos=0, final_scores=1.0255897,
final_local_scores=1.0526402, pem=0.0, similarity=35.502598]
18: 682925 Nasser_Hussain
    [pem_pos=3, final_scores=0.6891946,
final_local_scores=0.6914924, pem=-2.447519,
similarity=38.035213]
19: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.8432999,
final_local_scores=0.8691099, pem=-0.73930055,
similarity=34.046116]
20: 938861 Saleem_Malik
    [pem_pos=0, final_scores=0.7869471,
final_local_scores=0.82282126, pem=0.0, similarity=28.64434]
21: 3976606 Asif_Mujtaba
    [pem_pos=0, final_scores=0.6759338,
final_local_scores=0.7091225, pem=0.0, similarity=25.251343]
22: 501393 Wasim_Akram
```

```
[pem_pos=0, final_scores=0.7508696,
final_local_scores=0.74634916, pem=0.0, similarity=26.36226]
23: 1160680 Robert_Croft
    [pem_pos=7, final_scores=0.96448046,
final_local_scores=1.0009706, pem=-3.16984,
similarity=51.198757]
24: 554132 Moin_Khan
    [pem_pos=0, final_scores=0.65096074,
final_local_scores=0.6904409, pem=0.0, similarity=24.693846]
25: 896405 Mushtaq_Ahmed
    [pem_pos=0, final_scores=0.5555889,
final_local_scores=0.55059344, pem=0.0, similarity=20.52052]
26: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.71762383,
final_local_scores=0.73182786, pem=-0.73930055,
similarity=29.949348]
27: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.7508382,
final_local_scores=0.7502296, pem=0.0, similarity=26.47806]
28: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.69898874,
final_local_scores=0.71147186, pem=-0.73930055,
similarity=29.341885]
29: 1160680 Robert_Croft
    [pem_pos=7, final_scores=0.80216736,
final_local_scores=0.8236681, pem=-3.16984,
similarity=45.907696]
30: 2036840 Dominic_Cork
    [pem_pos=8, final_scores=0.9693317,
final_local_scores=0.98641807, pem=-3.606134,
similarity=53.13711]
31: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.71563053,
final_local_scores=0.68911606, pem=-3.3054674,
similarity=42.629955]
32: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.7665064,
final_local_scores=0.76342994, pem=0.0, similarity=26.871984]
33: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.7436637,
final_local_scores=0.7423926, pem=0.0, similarity=26.244188]
34: 896405 Mushtaq_Ahmed
    [pem_pos=0, final_scores=0.5902253,
final_local_scores=0.58842844, pem=0.0, similarity=21.649591]
```

Sample: b'330_CRICKET'

```
CRICKET - [1 1 PAKISTAN 1] DECLARE FIRST INNINGS AT 521-8 . [
2 2 LONDON 2] 1996-08-25 [3 3 Pakistan 3] declared their
first innings at 521-8 on the fourth day of the third and
final test against [4 4 England 4] at [5 5 The Oval 5] on
Sunday . Scores : [6 6 England 6] 326 ; [7 7 Pakistan 7]
521-8 .
```

```

1: 759295 Pakistan_national_cricket_team
2: 17867 London
3: 759295 Pakistan_national_cricket_team
4: 759125 England_cricket_team
5: 387687 The_Oval
6: 759125 England_cricket_team
7: 759295 Pakistan_national_cricket_team
1: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7903238,
final_local_scores=0.7610249, pem=-2.7556365,
similarity=41.785793]
2: 17867 London
    [pem_pos=0, final_scores=0.5491573,
final_local_scores=0.63170415, pem=-0.09086737,
similarity=23.435175]
3: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.83961123,
final_local_scores=0.81486386, pem=-2.7556365,
similarity=43.392456]
4: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.70114917,
final_local_scores=0.6607775, pem=-3.3054674,
similarity=41.78428]
5: 387687 The_Oval
    [pem_pos=0, final_scores=0.6183134,
final_local_scores=0.6108711, pem=0.0, similarity=22.319324]
6: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.71903557,
final_local_scores=0.6803157, pem=-3.3054674,
similarity=42.367336]
7: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7308492,
final_local_scores=0.69605786, pem=-2.7556365,
similarity=39.847046]

```

Sample: b'331_SOCCER'

```

SOCCER - [1 1 SCOTTISH 1] PREMIER DIVISION RESULT /
STANDINGS . [2 2 GLASGOW 2] 1996-08-25 Result of a [3 3 Scott
ish 3] premier division soccer match on Sunday : [4 4 Aberdeen
4] 4 [5 5 Hearts 5] 0 Standings ( tabulated under played ,
won , drawn , lost , goals for , against , points ) : [6 6 Ra
ngers 6] 3 3 0 0 7 2 9 [7 7 Celtic 7] 3 2 1 0 9 4 7 [8 8 Aber
deen 8] 3 1 2 0 8 4 5 [9 9 Motherwell 9] 3 1 2 0 6 3 5 [10 10
Hibernian 10] 3 1 1 1 2 2 4 [11 11 Hearts 11] 2 1 0 1 3 6 3 [1
2 12 Kilmarnock 12] 3 1 0 2 5 7 3 [13 13 Dundee United 13] 3
0 1 2 1 3 1 [14 14 Dunfermline 14] 2 0 1 1 2 5 1 [15 15 Raith
15] 3 0 0 3 1 8 0
1: 26994 Scotland
2: 68736 Glasgow
3: 26994 Scotland
4: 237019 Aberdeen_F.C.
5: 302212 Heart_of_Midlothian_F.C.

```

6: 69713 Rangers_F.C.
7: 69708 Celtic_F.C.
8: 237019 Aberdeen_F.C.
9: 215398 Motherwell_F.C.
10: 176747 Hibernian_F.C.
11: 302212 Heart_of_Midlothian_F.C.
12: 405663 Kilmarnock_F.C.
13: 331074 Dundee_United_F.C.
14: 451181 Dunfermline_Athletic_F.C.
15: 880348 Raith_Rovers_F.C.
1: 26994 Scotland
[pem_pos=0, final_scores=0.5145349,
final_local_scores=0.54279834, pem=-0.47951394,
similarity=22.895569]
2: 68736 Glasgow
[pem_pos=0, final_scores=0.5072863,
final_local_scores=0.52055633, pem=-0.15604196,
similarity=20.472733]
3: 26994 Scotland
[pem_pos=0, final_scores=0.5415446,
final_local_scores=0.5723024, pem=-0.47951394,
similarity=23.77603]
4: 237019 Aberdeen_F.C.
[pem_pos=1, final_scores=0.8342138,
final_local_scores=0.800671, pem=-1.4493513,
similarity=35.86513]
5: 302212 Heart_of_Midlothian_F.C.
[pem_pos=0, final_scores=0.62689143,
final_local_scores=0.5742094, pem=-0.29860955,
similarity=22.849152]
6: 69713 Rangers_F.C.
[pem_pos=0, final_scores=0.6752414,
final_local_scores=0.63225454, pem=-0.30191645,
similarity=24.599318]
7: 69708 Celtic_F.C.
[pem_pos=0, final_scores=0.6381742,
final_local_scores=0.59427696, pem=-0.9785444,
similarity=27.1456]
8: 237019 Aberdeen_F.C.
[pem_pos=1, final_scores=0.9207515,
final_local_scores=0.8952003, pem=-1.4493513,
similarity=38.686073]
9: 215398 Motherwell_F.C.
[pem_pos=0, final_scores=0.60850495,
final_local_scores=0.55666023, pem=-0.3361376,
similarity=22.529533]
10: 176747 Hibernian_F.C.
[pem_pos=0, final_scores=0.72918385,
final_local_scores=0.685139, pem=0.0, similarity=24.535627]
11: 302212 Heart_of_Midlothian_F.C.
[pem_pos=0, final_scores=0.6851418,
final_local_scores=0.6378391, pem=-0.29860955,

```
similarity=24.74799]
12: 405663 Kilmarnock_F.C.
    [pem_pos=1, final_scores=0.64714736,
final_local_scores=0.60409445, pem=-1.4326576,
similarity=29.90811]
13: 331074 Dundee_United_F.C.
    [pem_pos=0, final_scores=0.43277627,
final_local_scores=0.366853, pem=0.0, similarity=15.037338]
14: 451181 Dunfermline_Athletic_F.C.
    [pem_pos=1, final_scores=0.53711957,
final_local_scores=0.4829595, pem=-1.9918579,
similarity=29.33422]
15: 880348 Raith_Rovers_F.C.
    [pem_pos=0, final_scores=0.62241745,
final_local_scores=0.5838068, pem=0.0, similarity=21.511673]
```

Sample: b'332_CRICKET'
CRICKET - [1 1 PAKISTAN 1] 473-6 AT TEA ON FOURTH DAY THIRD TEST . [2 2 LONDON 2] 1996-08-25 [3 3 Pakistan 3] were 473-6 at tea on the fourth day of the third and final test at [4 4 The Oval 4] on Sunday in reply to [5 5 England 5] 's 326 . Scores : [6 6 England 6] 326 ; [7 7 Pakistan 7] 473-6 .

```
1: 759295 Pakistan_national_cricket_team
2: 17867 London
3: 759295 Pakistan_national_cricket_team
4: 387687 The_Oval
5: 759125 England_cricket_team
6: 759125 England_cricket_team
7: 759295 Pakistan_national_cricket_team
1: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.78466755,
final_local_scores=0.7548463, pem=-2.7556365,
similarity=41.60141]
2: 17867 London
    [pem_pos=0, final_scores=0.4965909,
final_local_scores=0.5742833, pem=-0.09086737,
similarity=21.721622]
3: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.8425663,
final_local_scores=0.8180919, pem=-2.7556365,
similarity=43.48879]
4: 387687 The_Oval
    [pem_pos=0, final_scores=0.6212643,
final_local_scores=0.6140945, pem=0.0, similarity=22.415518]
5: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.80790883,
final_local_scores=0.77739626, pem=-3.3054674,
similarity=45.264412]
6: 759125 England_cricket_team
    [pem_pos=3, final_scores=0.72402114,
final_local_scores=0.68576163, pem=-3.3054674,
similarity=42.529854]
```

```
7: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7786167,
final_local_scores=0.7482367, pem=-2.7556365,
similarity=41.404167]
```

Sample: b'333_CRICKET'

CRICKET - RUN-OUT GIVES LEWIS AND [1 1 ENGLAND 1] SLIM SATISFACTION . [2 2 LONDON 2] 1996-08-25 [3 3 Chris Lewis 3] did his best to forget his controversial omission from the [4 4 England 4] one-day squad on Sunday but could not prevent [5 5 Pakistan 5] reasserting their dominance in the final test at [6 6 The Oval 6] . A super piece of fielding by Lewis , dropped as a disciplinary measure after arriving only 35 minutes before the start on the fourth morning , provided the only bright spot for [7 7 England 7] as the touring team batted on to reach 413 for five at the interval , a lead of 87 . The solitary wicket to fall was [8 8 Asif Mujtaba 8] , run out for 13 attempting a second run to third man where Lewis was lurking with a point to prove . There seemed scant danger until the [9 9 Surrey 9] player swooped on to the ball and returned off balance to county team mate [10 10 Alec Stewart 10] who whipped off the bails . Lewis , though , could not make similar waves with the ball as [11 11 Salim Malik 11] and [12 12 Wasim Akram 12] batted through the rest of the session with few alarms . [13 13 Wasim 13] rattled along to 30 not out , outscoring his partner who was unbeaten on 24 at the break , although the weather was again threatening to play the dominant role . Rain arrived just as the players left the field for lunch , forcing the ground-staff into action yet again . The weather delayed the resumption for over an hour , the umpires finally announcing play would start again at 1445 local time (1345 [14 14 GMT 14]) .

```
1: 759125 England_cricket_team
2: 17867 London
3: 1732971 Chris_Lewis_(cricketer)
4: 759125 England_cricket_team
5: 759295 Pakistan_national_cricket_team
6: 387687 The_Oval
7: 759125 England_cricket_team
8: 3976606 Asif_Mujtaba
9: 766575 Surrey_County_Cricket_Club
10: 896414 Alec_Stewart
11: 938861 Saleem_Malik
12: 501393 Wasim_Akram
13: 501393 Wasim_Akram
14: 12701 Greenwich_Mean_Time
1: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.69658154,
final_local_scores=0.6573207, pem=-3.3054674,
similarity=41.681118]
2: 17867 London
    [pem_pos=0, final_scores=0.57433003,
```

```
final_local_scores=0.6708544, pem=-0.09086737,
similarity=24.603497]
3: 1732971 Chris_Lewis_(cricketer)
    [pem_pos=1, final_scores=0.54417086,
final_local_scores=0.5945524, pem=-0.7619004,
similarity=25.975676]
4: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.68845093,
final_local_scores=0.6484392, pem=-3.3054674,
similarity=41.416077]
5: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.8477989,
final_local_scores=0.8400124, pem=-2.7556365,
similarity=44.142937]
6: 387687 The_Oval
    [pem_pos=0, final_scores=0.5669883,
final_local_scores=0.55089426, pem=0.0, similarity=20.529497]
7: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.80558234,
final_local_scores=0.7763875, pem=-3.3054674,
similarity=45.23431]
8: 3976606 Asif_Mujtaba
    [pem_pos=0, final_scores=0.7625945,
final_local_scores=0.82602966, pem=0.0, similarity=28.740086]
9: 766575 Surrey_County_Cricket_Club
    [pem_pos=1, final_scores=0.5043572,
final_local_scores=0.49877292, pem=-1.7952642,
similarity=28.737019]
10: 896414 Alec_Stewart
    [pem_pos=0, final_scores=0.88105184,
final_local_scores=0.9136806, pem=0.0, similarity=31.355766]
11: 938861 Saleem_Malik
    [pem_pos=0, final_scores=0.78242856,
final_local_scores=0.82585007, pem=0.0, similarity=28.734726]
12: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.75978047,
final_local_scores=0.7581332, pem=0.0, similarity=26.713919]
13: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.8607356,
final_local_scores=0.8786964, pem=0.0, similarity=30.311766]
14: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.60438013,
final_local_scores=0.75911325, pem=0.0, similarity=26.743166]
```

Sample: b'334_CRICKET'

CRICKET - [1 1 ENGLAND 1] NAME SQUAD FOR ONE-DAY
INTERNATIONALS . [2 2 LONDON 2] 1996-08-25 The [3 3 England 3]
cricket squad was announced on Sunday for the one-day
international series against [4 4 Pakistan 4] starting on
Thursday . Squad : [5 5 Michael Atherton 5] (captain) , [6 6
Alec Stewart 6] , [7 7 Graham Thorpe 7] , [8 8 Nick Knight 8]
, [9 9 Graham Lloyd 9] , [10 10 Matthew Maynard 10] , [11 11 R

```
onnie_Irani_11] , [12_12 Adam_Hollioake_12] , [13_13 Robert_Croft_13] , [14_14 Darren_Gough_14] , [15_15 Peter_Martin_15] , [16_16 Dean_Headley_16] , [17_17 Alan_Mullally_17] .  
1: 759125 England_cricket_team  
2: 17867 London  
3: 9316 England  
4: 759295 Pakistan_national_cricket_team  
5: 159024 Michael_Atherton  
6: 896414 Alec_Stewart  
7: 397259 Graham_Thorpe  
8: 994718 Nick_Knight  
9: 3750666 Graham_Lloyd  
10: 2512623 Matthew_Maynard  
11: 2053953 Ronnie_Irani  
12: 2037175 Adam_Hollioake  
13: 1160680 Robert_Croft  
14: 413178 Darren_Gough  
15: 2062354 Peter_Martin_(cricketer)  
16: 672220 Dean_Headley  
17: 2222543 Alan_Mullally  
1: 759125 England_cricket_team  
    [pem_pos=2, final_scores=0.6797908,  
final_local_scores=0.6476148, pem=-3.3054674,  
similarity=41.391476]  
2: 17867 London  
    [pem_pos=0, final_scores=0.60454994,  
final_local_scores=0.7038644, pem=-0.09086737,  
similarity=25.588581]  
3: 9316 England  
    [pem_pos=0, final_scores=0.28010637,  
final_local_scores=0.3319857, pem=-0.19053856,  
similarity=15.033008]  
4: 759295 Pakistan_national_cricket_team  
    [pem_pos=1, final_scores=0.7969969,  
final_local_scores=0.812038, pem=-2.7556365,  
similarity=43.308125]  
5: 159024 Michael_Atherton  
    [pem_pos=0, final_scores=1.0389534,  
final_local_scores=1.0784254, pem=0.0, similarity=36.272076]  
6: 896414 Alec_Stewart  
    [pem_pos=0, final_scores=0.95940024,  
final_local_scores=0.9900618, pem=0.0, similarity=33.63513]  
7: 397259 Graham_Thorpe  
    [pem_pos=0, final_scores=0.7778984,  
final_local_scores=0.7949908, pem=0.0, similarity=27.813824]  
8: 994718 Nick_Knight  
    [pem_pos=1, final_scores=0.6242685,  
final_local_scores=0.67225665, pem=-0.55008924,  
similarity=27.142662]  
9: 3750666 Graham_Lloyd  
    [pem_pos=0, final_scores=0.5224554,  
final_local_scores=0.5822844, pem=0.0, similarity=21.46624]
```

```
10: 2512623 Matthew_Maynard
    [pem_pos=0, final_scores=0.6384662,
final_local_scores=0.6648541, pem=0.0, similarity=23.930286]
11: 2053953 Ronnie_Irani
    [pem_pos=0, final_scores=0.66436917,
final_local_scores=0.6912803, pem=0.0, similarity=24.718897]
12: 2037175 Adam_Hollioake
    [pem_pos=0, final_scores=0.57048875,
final_local_scores=0.6029977, pem=0.0, similarity=22.084368]
13: 1160680 Robert_Croft
    [pem_pos=0, final_scores=0.9985158,
final_local_scores=1.0536056, pem=0.0, similarity=35.531406]
14: 413178 Darren_Gough
    [pem_pos=0, final_scores=0.8883385,
final_local_scores=0.9252483, pem=0.0, similarity=31.700968]
15: 2062354 Peter_Martin_(cricketer)
    [pem_pos=1, final_scores=0.5391451,
final_local_scores=0.58913296, pem=-1.2324854,
similarity=28.373062]
16: 672220 Dean_Headley
    [pem_pos=0, final_scores=0.49145532,
final_local_scores=0.55177814, pem=0.0, similarity=20.555874]
17: 2222543 Alan_Mullally
    [pem_pos=0, final_scores=0.65393364,
final_local_scores=0.67573243, pem=0.0, similarity=24.254917]
```

Sample: b'335_TENNIS'

TENNIS - RESULTS AT [1 1 CANADIAN OPEN 1] . [2 2 TORONTO 2]
1996-08-25 Results from the [3 3 Canadian Open 3] tennis
tournament on Sunday (prefix number denotes seeding) :
Final 3 - [4 4 Wayne Ferreira 4] ([5 5 South Africa 5])
beat [6 6 Todd Woodbridge 6] ([7 7 Australia 7]) 6-2 6-4

```
1: 1872219 Canadian_Open_(tennis)
2: 64646 Toronto
3: 1872219 Canadian_Open_(tennis)
4: 1723516 Wayne_Ferreira
5: 17416221 South_Africa
6: 1547951 Todd_Woodbridge
7: 4689264 Australia
1: 1872219 Canadian_Open_(tennis)
    [pem_pos=0, final_scores=0.5734939,
final_local_scores=0.5798668, pem=-0.58982956,
similarity=24.60168]
2: 64646 Toronto
    [pem_pos=0, final_scores=0.5398968,
final_local_scores=0.61353713, pem=-0.12291032,
similarity=23.067291]
3: 1872219 Canadian_Open_(tennis)
    [pem_pos=0, final_scores=0.47837025,
final_local_scores=0.47595865, pem=-0.58982956,
similarity=21.500853]
4: 1723516 Wayne_Ferreira
```

```
[pem_pos=0, final_scores=0.8109402,
final_local_scores=0.9132039, pem=0.0, similarity=31.341537]
5: 17416221 South_Africa
[pem_pos=0, final_scores=0.54218847,
final_local_scores=0.6457435, pem=-0.26251382,
similarity=24.787579]
6: 1547951 Todd_Woodbridge
[pem_pos=0, final_scores=0.6913708,
final_local_scores=0.7823663, pem=0.0, similarity=27.437082]
7: 4689264 Australia
[pem_pos=0, final_scores=0.6559609,
final_local_scores=0.7534846, pem=-0.17168477,
similarity=27.508846]
```

Sample: b'336_TENNIS'

TENNIS - RESULTS AT [1 1 CANADIAN OPEN 1] . [2 2 TORONTO 2]
1996-08-24 Results from the [3 3 Canadian Open 3] tennis
tournament on Saturday (prefix numbers denotes seedings) :
Semifinals 3 - [4 4 Wayne Ferreira 4] ([5 5 South Africa
5]) beat 7 - [6 6 Todd Martin 6] ([7 7 U.S. 7]) 4-6 6-3
7-5 [8 8 Todd Woodbridge 8] ([9 9 Australia 9]) beat 4 - [10
10 Marcelo Rios 10] ([11 11 Chile 11]) 6-0 6-3

```
1: 1872219 Canadian_Open_(tennis)
2: 64646 Toronto
3: 1872219 Canadian_Open_(tennis)
4: 1723516 Wayne_Ferreira
5: 17416221 South_Africa
6: 871725 Todd_Martin
7: 3434750 United_States
8: 1547951 Todd_Woodbridge
9: 4689264 Australia
10: 483554 Marcelo_Ríos
11: 5489 Chile
1: 1872219 Canadian_Open_(tennis)
[pem_pos=0, final_scores=0.547782,
final_local_scores=0.55880976, pem=-0.58982956,
similarity=23.973295]
2: 64646 Toronto
[pem_pos=0, final_scores=0.56019205,
final_local_scores=0.61555713, pem=-0.12291032,
similarity=23.127571]
3: 1872219 Canadian_Open_(tennis)
[pem_pos=0, final_scores=0.46570092,
final_local_scores=0.46914864, pem=-0.58982956,
similarity=21.297628]
4: 1723516 Wayne_Ferreira
[pem_pos=0, final_scores=0.83705026,
final_local_scores=0.93567365, pem=0.0, similarity=32.01208]
5: 17416221 South_Africa
[pem_pos=0, final_scores=0.60389024,
final_local_scores=0.68857414, pem=-0.26251382,
similarity=26.06573]
```

```
6: 871725 Todd_Martin
    [pem_pos=0, final_scores=0.8803721,
final_local_scores=0.9479447, pem=0.0, similarity=32.378273]
7: 3434750 United_States
    [pem_pos=0, final_scores=0.4798879,
final_local_scores=0.52777964, pem=-0.14075686,
similarity=20.605167]
8: 1547951 Todd_Woodbridge
    [pem_pos=0, final_scores=0.84333116,
final_local_scores=0.93775195, pem=0.0, similarity=32.0741]
9: 4689264 Australia
    [pem_pos=0, final_scores=0.62771636,
final_local_scores=0.6922558, pem=-0.17168477,
similarity=25.681656]
10: 483554 Marcelo_Ríos
    [pem_pos=0, final_scores=0.69227487,
final_local_scores=0.7357777, pem=0.0, similarity=26.046787]
11: 5489 Chile
    [pem_pos=0, final_scores=0.5578334,
final_local_scores=0.63228613, pem=-0.27404457,
similarity=24.448689]
```

Sample: b'337_SOCCER'

```
SOCCER - [1 1 TOGO 1] BEAT [2 2 CONGO 2] 1-0 IN [3 3 AFRICA NATIONS CUP 3] QUALIFIER . [4 4 LOME 4] 1996-08-25 [5 5 Togo 5] beat [6 6 Congo 6] 1-0 ( halftime 0-0 ) in their [7 7 African Nations Cup 7] preliminary round , second leg qualifying match on Sunday . Scorer : Salou Bachirou ( 53rd minute ) Attendance : 18,000 [8 8 Togo 8] win 1-0 on aggregate .
```

```
1: 1080720 Togo_national_football_team
2: 1047784 Congo_national_football_team
3: 250680 Africa_Cup_of_Nations
4: 57067 Lomé
5: 1080720 Togo_national_football_team
6: 1047784 Congo_national_football_team
7: 250680 Africa_Cup_of_Nations
8: 1080720 Togo_national_football_team
1: 1080720 Togo_national_football_team
    [pem_pos=1, final_scores=0.6589867,
final_local_scores=0.6311851, pem=-2.2790594,
similarity=35.319416]
2: 1047784 Congo_national_football_team
    [pem_pos=2, final_scores=0.61881155,
final_local_scores=0.61996657, pem=-2.414126,
similarity=35.719143]
3: 250680 Africa_Cup_of_Nations
    [pem_pos=0, final_scores=0.5947342,
final_local_scores=0.5742916, pem=0.0, similarity=21.22772]
4: 57067 Lomé
    [pem_pos=0, final_scores=0.41653538,
final_local_scores=0.4658612, pem=0.0, similarity=17.99194]
5: 1080720 Togo_national_football_team
```

```
[pem_pos=1, final_scores=0.53331375,
final_local_scores=0.49390656, pem=-2.2790594,
similarity=31.222748]
6: 1047784 Congo_national_football_team
    [pem_pos=2, final_scores=0.90517586,
final_local_scores=0.93277615, pem=-2.414126,
similarity=45.054005]
7: 250680 Africa_Cup_of_Nations
    [pem_pos=0, final_scores=0.5614683,
final_local_scores=0.5379537, pem=0.0, similarity=20.143324]
8: 1080720 Togo_national_football_team
    [pem_pos=1, final_scores=0.68177426,
final_local_scores=0.6560771, pem=-2.2790594,
similarity=36.06224]
```

Sample: b'338_SOCCER'
SOCCER - [1 1 ETHIOPIA 1] BEAT [2 2 UGANDA 2] ON PENALTIES
IN [3 3 AFRICAN NATIONS CUP 3] . ADDIS ABABA 1996-08-25 [4 4 E
thiopia 4] and [5 5 Uganda 5] drew 1-1 (halftime 1-0) in
their [6 6 African Nations Cup 6] preliminary round , second
leg match on Sunday . Attendance : 25,000 Aggregate 2-2 . [7 7
Ethiopia 7] won 4-2 penalties .

```
1: 1089708 Ethiopia_national_football_team
2: 1092055 Uganda_national_football_team
3: 250680 Africa_Cup_of_Nations
4: 1089708 Ethiopia_national_football_team
5: 1092055 Uganda_national_football_team
6: 250680 Africa_Cup_of_Nations
7: 1089708 Ethiopia_national_football_team
1: 1089708 Ethiopia_national_football_team
    [pem_pos=1, final_scores=0.5313632,
final_local_scores=0.48470807, pem=-2.7670166,
similarity=33.601833]
2: 1092055 Uganda_national_football_team
    [pem_pos=1, final_scores=0.5105943,
final_local_scores=0.4848274, pem=-2.80235,
similarity=33.797543]
3: 250680 Africa_Cup_of_Nations
    [pem_pos=0, final_scores=0.60609984,
final_local_scores=0.58437246, pem=0.0, similarity=21.528553]
4: 1089708 Ethiopia_national_football_team
    [pem_pos=1, final_scores=0.6398217,
final_local_scores=0.60318255, pem=-2.7670166,
similarity=37.13735]
5: 1092055 Uganda_national_football_team
    [pem_pos=1, final_scores=0.68432474,
final_local_scores=0.67460155, pem=-2.80235,
similarity=39.460785]
6: 250680 Africa_Cup_of_Nations
    [pem_pos=0, final_scores=0.5851803,
final_local_scores=0.561521, pem=0.0, similarity=20.846619]
7: 1089708 Ethiopia_national_football_team
```

```
[pem_pos=1, final_scores=0.6904802,
final_local_scores=0.65851927, pem=-2.7670166,
similarity=38.788708]

Sample: b'339_SOCCER'
SOCCER - [1 1 YUGOSLAV 1] LEAGUE RESULTS . [2 2 BELGRADE 2]
1996-08-25 Results of [3 3 Yugoslav 3] league soccer matches
played on Sunday : Division A [4 4 Vojvodina 4] 1 [5 5 Partizan 5] 1 [6 6 Crvena zvezda 6] 3 [7 7 Proleter 7] 1 Division B
Zelesnik 0 [8 8 Rudar 8] 1
1: 34244 Yugoslavia
2: 55904 Belgrade
3: 34244 Yugoslavia
4: 2652716 FK_Vojvodina
5: 381658 FK_Partizan
6: 361966 Red_Star_Belgrade
7: 12576085 FK_Proleter_Zrenjanin
8: 6885117 FK_Rudar_Pljevlja
1: 34244 Yugoslavia
[pem_pos=0, final_scores=0.29902035,
final_local_scores=0.28288502, pem=-0.7596219,
similarity=16.662508]
2: 55904 Belgrade
[pem_pos=0, final_scores=0.47084987,
final_local_scores=0.4577604, pem=-0.05193398,
similarity=18.032621]
3: 34244 Yugoslavia
[pem_pos=0, final_scores=0.3059777,
final_local_scores=0.2904849, pem=-0.7596219,
similarity=16.889303]
4: 2652716 FK_Vojvodina
[pem_pos=1, final_scores=0.33891118,
final_local_scores=0.29711574, pem=-1.2114366,
similarity=19.54422]
5: 381658 FK_Partizan
[pem_pos=0, final_scores=0.32956225,
final_local_scores=0.28353906, pem=-0.1934529,
similarity=13.60311]
6: 361966 Red_Star_Belgrade
[pem_pos=0, final_scores=0.47346544,
final_local_scores=0.43266928, pem=0.0, similarity=17.001427]
7: 12576085 FK_Proleter_Zrenjanin
[pem_pos=0, final_scores=0.28595012,
final_local_scores=0.28669223, pem=0.0, similarity=12.645181]
8: 6885117 FK_Rudar_Pljevlja
[pem_pos=0, final_scores=0.18934105,
final_local_scores=0.1913107, pem=-0.37007397,
similarity=11.811326]
```

```
Sample: b'340_SOCCER'
SOCCER - LEADING GOALSCORERS IN [1 1 POLISH 1] FIRST
DIVISION . [2 2 WARSAW 2] 1996-08-25 Leading goalscorers in
```

the [3 3 Polish 3] first division after the weekend 's matches : 7 - Bogdan Prusek (Sokol Tychy) 5 - [4 4 Sławomir Wojciechowski 4] ([5 5 GKS Katowice 5]) 4 - [6 6 Jacek Dembiński 6] ([7 7 Widzew Łódź 7]) , [8 8 Marcin Mięciel 8] ([9 9 Legia Warsaw 9]) , [10 10 Ryszard Wieczorek 10] ([11 11 Odra Wodzisław 11]) 3 - Jacek Berensztain ([12 12 GKS Bełchatów 12]) , [13 13 Marek Citko 13] ([14 14 Widzew 14]) , [15 15 Adam Fedoruk 15] , Dariusz Jackiewicz (both [16 16 Amica Wroneki 16]) , Bartłomiej Jamroz ([17 17 Hutnik Kraków 17]) , [18 18 Tomasz Moskal 18] ([19 19 Śląsk Wrocław 19]) , [20 20 Krzysztof Piskuła 20] ([21 21 Lech Poznań 21]) , Mariusz Srutwa ([22 22 Ruch Chorzów 22]) , [23 23 Emmanuel Tetteh 23] ([24 24 Polonia Warszawa 24]) , Krzysztof Zagórski ([25 25 Odra 25])

1: 22936 Poland
2: 32908 Warsaw
3: 22936 Poland
4: 21327266 Sławomir_Wojciechowski
5: 677596 GKS_Katowice
6: 21381761 Jacek_Dembiński
7: 677613 Widzew_Łódź
8: 8424753 Marcin_Mięciel
9: 677536 Legia_Warsaw
10: 8581102 Ryszard_Wieczorek
11: 677572 Odra_Wodzisław
12: 677651 GKS_Bełchatów
13: 3000120 Marek_Citko
14: 677613 Widzew_Łódź
15: 21075419 Adam_Fedoruk
16: 677540 Amica_Wroneki
17: 805862 Hutnik_Nowa_Huta
18: 13122882 Tomasz_Moskal
19: 772146 Śląsk_Wrocław
20: 19739035 Krzysztof_Piskuła
21: 677622 Lech_Poznań
22: 677685 Ruch_Chorzów
23: 23958681 Emmanuel_Tetteh
24: 677603 Polonia_Warsaw
25: 677572 Odra_Wodzisław

1: 22936 Poland
[pem_pos=0, final_scores=0.57132494, final_local_scores=0.62658566, pem=-0.81518245, similarity=27.221367]
2: 32908 Warsaw
[pem_pos=0, final_scores=0.4904458, final_local_scores=0.56076145, pem=-0.14300999, similarity=21.601664]
3: 22936 Poland
[pem_pos=0, final_scores=0.5206162, final_local_scores=0.571194, pem=-0.81518245, similarity=25.56837]
4: 21327266 Sławomir_Wojciechowski

[pem_pos=0, final_scores=0.3806042,
final_local_scores=0.49249452, pem=0.0, similarity=18.786732]
5: 677596 GKS_Katowice
[pem_pos=0, final_scores=0.5463496,
final_local_scores=0.5367156, pem=0.0, similarity=20.106379]
6: 21381761 Jacek_Dembiński
[pem_pos=0, final_scores=0.5513578,
final_local_scores=0.6693045, pem=0.0, similarity=24.063093]
7: 677613 Widzew_Łódź
[pem_pos=0, final_scores=0.5379966,
final_local_scores=0.50925744, pem=0.0, similarity=19.28697]
8: 8424753 Marcin_Mięciel
[pem_pos=0, final_scores=0.8133294,
final_local_scores=0.91236323, pem=0.0, similarity=31.316452]
9: 677536 Legia_Warsaw
[pem_pos=0, final_scores=0.50774544,
final_local_scores=0.4744901, pem=0.0, similarity=18.249445]
10: 8581102 Ryszard_Wieczorek
[pem_pos=0, final_scores=0.47414273,
final_local_scores=0.5931453, pem=0.0, similarity=21.790352]
11: 677572 Odra_Wodzisław
[pem_pos=0, final_scores=0.66522944,
final_local_scores=0.680827, pem=0.0, similarity=24.40695]
12: 677651 GKS_Bełchatów
[pem_pos=0, final_scores=0.6092047,
final_local_scores=0.6250883, pem=0.0, similarity=22.743593]
13: 3000120 Marek_Citko
[pem_pos=0, final_scores=0.856962,
final_local_scores=0.9462485, pem=0.0, similarity=32.327656]
14: 677613 Widzew_Łódź
[pem_pos=0, final_scores=0.52454513,
final_local_scores=0.4945637, pem=0.0, similarity=18.84848]
15: 21075419 Adam_Fedoruk
[pem_pos=0, final_scores=0.49151856,
final_local_scores=0.5991349, pem=0.0, similarity=21.969093]
16: 677540 Amica_Wronki
[pem_pos=0, final_scores=0.88402945,
final_local_scores=0.9199677, pem=0.0, similarity=31.543383]
17: 805862 Hutnik_Nowa_Huta
[pem_pos=0, final_scores=0.53228563,
final_local_scores=0.5709005, pem=0.0, similarity=21.126522]
18: 13122882 Tomasz_Moskal
[pem_pos=0, final_scores=0.23785006,
final_local_scores=0.44543403, pem=0.0, similarity=17.382353]
19: 772146 Śląsk_Wrocław
[pem_pos=0, final_scores=0.50769264,
final_local_scores=0.49032617, pem=0.0, similarity=18.722023]
20: 19739035 Krzysztof_Piskuła
[pem_pos=0, final_scores=0.5036163,
final_local_scores=0.63185674, pem=0.0, similarity=22.945578]
21: 677622 Lech_Poznań
[pem_pos=0, final_scores=0.56280714,

```

final_local_scores=0.54838365, pem=0.0, similarity=20.454575]
22: 677685 Ruch_Chorzów
    [pem_pos=0, final_scores=0.6398459,
final_local_scores=0.6326033, pem=0.0, similarity=22.967857]
23: 23958681 Emmanuel_Tetteh
    [pem_pos=0, final_scores=0.5674968,
final_local_scores=0.6993665, pem=0.0, similarity=24.960205]
24: 677603 Polonia_Warsaw
    [pem_pos=0, final_scores=0.5260939,
final_local_scores=0.49575394, pem=0.0, similarity=18.883999]
25: 677572 Odra_Wodzisław
    [pem_pos=4, final_scores=0.40536338,
final_local_scores=0.39696282, pem=-1.5730703,
similarity=24.490477]

```

Sample: b'341_SOCCER'

```

SOCCER - [1 1 ROMANIAN 1] FIRST DIVISION RESULTS /
STANDINGS . [2 2 BUCHAREST 2] 1996-08-25 Results of first
division soccer matches played over the weekend : A.S. Bacau
1 [3 3 Ceahlaul Piatra Neamt 3] 1 [4 4 Otelul Galati 4] 1
F.C. Arges Dacia Pitesti 0 F.C. Farul Constanta 3 Chindia
Tirgoviste 1 [5 5 Sportul Studentesc 5] 4 [6 6 Universitatea C
raiova 6] 2 F.C. Petrolul Ploiesti 4 Politehnica Timisoara 5
F.C. Brasov 1 F.C. National Bucharest 1 [7 7 Jiul Petrosani 7]
1 [8 8 Dinamo Bucharest 8] 0 [9 9 Gloria Bistrita 9] 0 [10 10
Universitatea Cluj 10] 1 [11 11 Rapid Bucharest 11] 0 [12 12 S
teaua Bucharest 12] 2 Standings ( tabulate under played ,
won , drawn , lost , goals for , against , points ) : [13 13 D
inamo Bucharest 13] 4 3 0 1 6 2 9 [14 14 Jiul Petrosani 14] 4
3 0 1 6 4 9 F.C. Farul Constanta 4 2 2 0 6 2 8 [15 15 Univers
itatea Cluj 15] 4 2 2 0 6 4 8 A.S. Bacau 4 2 1 1 7 3 7
Politehnica Timisoara 4 2 1 1 11 9 7 F.C. National Bucharest
4 2 1 1 7 6 7 [16 16 Otelul Galati 16] 4 2 0 2 4 3 6 F.C.
Chindia Tirgoviste 4 2 0 2 3 4 6 [17 17 Steaua Bucharest 17]
4 2 0 2 5 7 6 F.C. Arges Dacia Pitesti 4 1 2 1 4 2 5 [18 18 Un
iversitatea Craiova 18] 4 1 1 2 8 6 4 [19 19 Sportul Studente
sc 19] 4 1 1 2 7 9 4 [20 20 Ceahlaul Piatra Neamt 20] 4 1 1 2
2 4 4 F.C. Brasov 4 1 1 2 6 10 4 [21 21 Gloria Bistrita 21] 4
1 0 3 3 9 3 F.C. Petrolul Ploiesti 4 0 2 2 5 7 2 [22 22 Rapid
Bucharest 22] 4 0 1 3 4 9 1

```

- 1: 25445 Romania
- 2: 36877 Bucharest
- 3: 2958305 FC_Ceahlăul_Piatra_Neamț
- 4: 3569554 FC_Oțelul_Galați
- 5: 2498018 FC_Sportul_Studențesc_București
- 6: 2384859 FC_Universitatea_Craiova
- 7: 3574206 CS_Jiul_Petroșani
- 8: 861358 FC_Dinamo_București
- 9: 2384775 ACF_Gloria_Bistrița
- 10: 2554958 FC_Universitatea_Cluj
- 11: 1240318 FC_Rapid_București
- 12: 893411 FC_Steaua_București

13: 861358 FC_Dinamo_Bucureşti
14: 3574206 CS_Jiul_Petroşani
15: 2554958 FC_Universitatea_Cluj
16: 3569554 FC_Otelul_Galaţi
17: 893411 FC_Steaua_Bucureşti
18: 2384859 FC_Universitatea_Craiova
19: 2498018 FC_Sportul_Studenţesc_Bucureşti
20: 2958305 FC_Ceahlăul_Piatra_Neamţ
21: 2384775 ACF_Gloria_Bistriţa
22: 1240318 FC_Rapid_Bucureşti
1: 25445 Romania
 [pem_pos=1, final_scores=0.64410895,
final_local_scores=0.7155062, pem=-1.0382323,
similarity=31.087914]
2: 36877 Bucharest
 [pem_pos=0, final_scores=0.4581355,
final_local_scores=0.49672192, pem=-0.01781679,
similarity=19.009777]
3: 2958305 FC_Ceahlăul_Piatra_Neamţ
 [pem_pos=0, final_scores=0.4340037,
final_local_scores=0.4157101, pem=0.0, similarity=16.49533]
4: 3569554 FC_Otelul_Galaţi
 [pem_pos=0, final_scores=0.54313713,
final_local_scores=0.51232576, pem=0.0, similarity=19.378536]
5: 2498018 FC_Sportul_Studenţesc_Bucureşti
 [pem_pos=0, final_scores=0.4601168,
final_local_scores=0.42960352, pem=0.0, similarity=16.909939]
6: 2384859 FC_Universitatea_Craiova
 [pem_pos=0, final_scores=0.41716927,
final_local_scores=0.37486023, pem=0.0, similarity=15.276289]
7: 3574206 CS_Jiul_Petroşani
 [pem_pos=0, final_scores=0.43799305,
final_local_scores=0.44326633, pem=0.0, similarity=17.317665]
8: 861358 FC_Dinamo_Bucureşti
 [pem_pos=0, final_scores=0.34615332,
final_local_scores=0.28142223, pem=0.0, similarity=12.487913]
9: 2384775 ACF_Gloria_Bistriţa
 [pem_pos=0, final_scores=0.6713832,
final_local_scores=0.65544057, pem=0.0, similarity=23.649366]
10: 2554958 FC_Universitatea_Cluj
 [pem_pos=0, final_scores=0.3686232,
final_local_scores=0.3170337, pem=0.0, similarity=13.5506315]
11: 1240318 FC_Rapid_Bucureşti
 [pem_pos=0, final_scores=0.49797934,
final_local_scores=0.44715875, pem=0.0, similarity=17.433823]
12: 893411 FC_Steaua_Bucureşti
 [pem_pos=1, final_scores=0.41286737,
final_local_scores=0.3501679, pem=-0.074108,
similarity=14.942433]
13: 861358 FC_Dinamo_Bucureşti
 [pem_pos=0, final_scores=0.41713053,
final_local_scores=0.35895404, pem=0.0, similarity=14.801618]

```

14: 3574206 CS_Jiul_Petroșani
    [pem_pos=0, final_scores=0.43587202,
final_local_scores=0.44094944, pem=0.0, similarity=17.248524]
15: 2554958 FC_Universitatea_Cluj
    [pem_pos=0, final_scores=0.36665338,
final_local_scores=0.31488192, pem=0.0, similarity=13.486418]
16: 3569554 FC_otelul_Galați
    [pem_pos=0, final_scores=0.5317924,
final_local_scores=0.49993336, pem=0.0, similarity=19.008722]
17: 893411 FC_Steaua_București
    [pem_pos=1, final_scores=0.40938705,
final_local_scores=0.34636623, pem=-0.074108,
similarity=14.828983]
18: 2384859 FC_Universitatea_Craiova
    [pem_pos=0, final_scores=0.46091098,
final_local_scores=0.42264146, pem=0.0, similarity=16.702177]
19: 2498018 FC_Sportul_Studentesc_București
    [pem_pos=0, final_scores=0.45866746,
final_local_scores=0.42802036, pem=0.0, similarity=16.862694]
20: 2958305 FC_Ceahlăul_Piatra_Neamț
    [pem_pos=0, final_scores=0.44705015,
final_local_scores=0.42996138, pem=0.0, similarity=16.920618]
21: 2384775 ACF_Gloria_Bistrița
    [pem_pos=0, final_scores=0.70752245,
final_local_scores=0.69491726, pem=0.0, similarity=24.82743]
22: 1240318 FC_Rapid_București
    [pem_pos=0, final_scores=0.4532811,
final_local_scores=0.39833272, pem=0.0, similarity=15.976756]

```

Sample: b'342_SOCCER'

SOCCER - [1 1 POLISH 1] FIRST DIVISION RESULTS / STANDINGS .
[2 2 WARSAW 2] 1996-08-25 Results of [3 3 Polish 3] first
division soccer matches played over the weekend : [4 4 Amica W
ronki 4] 3 [5 5 Hutnik Krakow 5] 0 Sokol Tychy 5 [6 6 Lech Po
znan 6] 3 [7 7 Rakow Czestochowa 7] 1 [8 8 Stomil Olsztyn 8]
4 [9 9 Wisla Krakow 9] 1 [10 10 Gornik Zabrze 10] 0 [11 11 Sl
ask Wroclaw 11] 3 [12 12 Odra Wodzislaw 12] 1 [13 13 GKS Kato
wice 13] 1 [14 14 Polonia Warsaw 14] 0 [15 15 Zaglebie Lubin 1
5] 2 [16 16 LKS Lodz 16] 1 [17 17 Legia Warsaw 17] 3 [18 18 GK
S Belchatow 18] 2 [19 19 Widzew Lodz 19] 3 [20 20 Ruch Chorzow
20] 0 Standings (tabulate under played , won , drawn ,
lost , goals for , against , points) : [21 21 Amica Wronki 21
] 7 5 1 1 13 8 16 [22 22 Legia Warsaw 22] 7 5 1 1 13 7 16 [23
23 Lech Poznan 23] 7 5 0 2 12 9 15 [24 24 Widzew Lodz 24] 7 4
2 1 13 3 14 [25 25 GKS Katowice 25] 7 4 2 1 12 9 14 Sokol
Tychy 7 4 0 3 14 15 12 [26 26 Odra Wodzislaw 26] 7 3 1 3 13
10 10 [27 27 Slask Wroclaw 27] 7 3 1 3 8 7 10 [28 28 Polonia W
arsaw 28] 7 3 1 3 7 9 10 [29 29 GKS Belchatow 29] 7 3 0 4 9 9
9 [30 30 Stomil Olsztyn 30] 7 2 3 2 9 9 9 [31 31 Wisla Krakow
31] 7 2 3 2 3 4 9 [32 32 Hutnik Krakow 32] 7 3 0 4 8 10 9 [33
33 Rakow Czestochowa 33] 7 2 1 4 6 10 7 [34 34 Zaglebie Lubin
34] 7 1 3 3 10 12 6 [35 35 Ruch Chorzow 35] 7 1 2 4 7 13 5 [36

36 Gornik Zabrze 36] 7 1 1 5 6 10 4 [37 37 LKS Lodz 37] 7 0 2
5 4 13 2
1: 22936 Poland
2: 32908 Warsaw
3: 22936 Poland
4: 677540 Amica_Wronki
5: 805862 Hutnik_Nowa_Huta
6: 677622 Lech_Poznań
7: 8538149 Raków_Częstochowa
8: 10410246 Stomil_Olsztyn_(football)
9: 677524 Wisła_Kraków
10: 677580 Górnik_Zabrze
11: 772146 Śląsk_Wrocław
12: 677572 Odra_Wodzisław
13: 677596 GKS_Katowice
14: 677603 Polonia_Warsaw
15: 677663 Zagłębie_Lubin
16: 677683 ŁKS_Łódź
17: 677536 Legia_Warsaw
18: 677651 GKS_Bełchatów
19: 677613 Widzew_Łódź
20: 677685 Ruch_Chorzów
21: 677540 Amica_Wronki
22: 677536 Legia_Warsaw
23: 677622 Lech_Poznań
24: 677613 Widzew_Łódź
25: 677596 GKS_Katowice
26: 677572 Odra_Wodzisław
27: 772146 Śląsk_Wrocław
28: 677603 Polonia_Warsaw
29: 677651 GKS_Bełchatów
30: 10410246 Stomil_Olsztyn_(football)
31: 677524 Wisła_Kraków
32: 805862 Hutnik_Nowa_Huta
33: 8538149 Raków_Częstochowa
34: 677663 Zagłębie_Lubin
35: 677685 Ruch_Chorzów
36: 677580 Górnik_Zabrze
37: 677683 ŁKS_Łódź
1: 22936 Poland
[pem_pos=0, final_scores=0.60633886,
final_local_scores=0.6757594, pem=-0.81518245,
similarity=28.688808]
2: 32908 Warsaw
[pem_pos=0, final_scores=0.46826512,
final_local_scores=0.54485697, pem=-0.14300999,
similarity=21.127043]
3: 22936 Poland
[pem_pos=0, final_scores=0.54984254,
final_local_scores=0.6140457, pem=-0.81518245,
similarity=26.847149]
4: 677540 Amica_Wronki

[pem_pos=0, final_scores=0.66832596,
final_local_scores=0.6710784, pem=0.0, similarity=24.11603]
5: 805862 Hutnik_Nowa_Huta
[pem_pos=0, final_scores=0.5368957,
final_local_scores=0.5613611, pem=0.0, similarity=20.841846]
6: 677622 Lech_Poznań
[pem_pos=0, final_scores=0.6112885,
final_local_scores=0.5883942, pem=0.0, similarity=21.648571]
7: 8538149 Raków_Częstochowa
[pem_pos=0, final_scores=0.5596203,
final_local_scores=0.5841832, pem=0.0, similarity=21.522905]
8: 10410246 Stomil_Olsztyn_(football)
[pem_pos=0, final_scores=0.52857536,
final_local_scores=0.54216594, pem=0.0, similarity=20.269026]
9: 677524 Wisła_Kraków
[pem_pos=0, final_scores=0.60928345,
final_local_scores=0.5713455, pem=0.0, similarity=21.139803]
10: 677580 Górnik_Zabrze
[pem_pos=0, final_scores=0.65115416,
final_local_scores=0.61793995, pem=0.0, similarity=22.530273]
11: 772146 Śląsk_Wrocław
[pem_pos=0, final_scores=0.52993363,
final_local_scores=0.5014365, pem=0.0, similarity=19.053577]
12: 677572 Odra_Wodzisław
[pem_pos=0, final_scores=0.6439021,
final_local_scores=0.6542266, pem=0.0, similarity=23.61314]
13: 677596 GKS_Katowice
[pem_pos=0, final_scores=0.5555894,
final_local_scores=0.5316397, pem=0.0, similarity=19.954903]
14: 677603 Polonia_Warsaw
[pem_pos=0, final_scores=0.5839414,
final_local_scores=0.549049, pem=0.0, similarity=20.47443]
15: 677663 Zagłębie_Lubin
[pem_pos=0, final_scores=0.6434754,
final_local_scores=0.62441814, pem=0.0, similarity=22.723597]
16: 677683 ŁKS_Łódź
[pem_pos=0, final_scores=0.559261,
final_local_scores=0.52286714, pem=0.0, similarity=19.693111]
17: 677536 Legia_Warsaw
[pem_pos=0, final_scores=0.5088982,
final_local_scores=0.46570092, pem=0.0, similarity=17.987158]
18: 677651 GKS_Bełchatów
[pem_pos=0, final_scores=0.5713411,
final_local_scores=0.5705822, pem=0.0, similarity=21.117023]
19: 677613 Widzew_Łódź
[pem_pos=0, final_scores=0.49042207,
final_local_scores=0.4490348, pem=0.0, similarity=17.489807]
20: 677685 Ruch_Chorzów
[pem_pos=0, final_scores=0.6721186,
final_local_scores=0.6539117, pem=0.0, similarity=23.603743]
21: 677540 Amica_Wronki
[pem_pos=0, final_scores=0.692326,

```
final_local_scores=0.69729483, pem=0.0, similarity=24.898382]
22: 677536 Legia_Warsaw
    [pem_pos=0, final_scores=0.48085648,
final_local_scores=0.43506962, pem=0.0, similarity=17.07306]
23: 677622 Lech_Poznań
    [pem_pos=0, final_scores=0.5473342,
final_local_scores=0.51853395, pem=0.0, similarity=19.5638]
24: 677613 Widzew_Łódź
    [pem_pos=0, final_scores=0.45116407,
final_local_scores=0.40615147, pem=0.0, similarity=16.210083]
25: 677596 GKS_Katowice
    [pem_pos=0, final_scores=0.5157033,
final_local_scores=0.4880702, pem=0.0, similarity=18.654701]
26: 677572 Odra_Wodzisław
    [pem_pos=0, final_scores=0.6252678,
final_local_scores=0.63387144, pem=0.0, similarity=23.005701]
27: 772146 Śląsk_Wrocław
    [pem_pos=0, final_scores=0.49740493,
final_local_scores=0.46590376, pem=0.0, similarity=17.99321]
28: 677603 Polonia_Warsaw
    [pem_pos=0, final_scores=0.53521687,
final_local_scores=0.4958248, pem=0.0, similarity=18.886114]
29: 677651 GKS_Bełchatów
    [pem_pos=0, final_scores=0.5523239,
final_local_scores=0.54980874, pem=0.0, similarity=20.497103]
30: 10410246 Stomil_Olsztyn_(football)
    [pem_pos=0, final_scores=0.476412,
final_local_scores=0.48518533, pem=0.0, similarity=18.568611]
31: 677524 Wisła_Kraków
    [pem_pos=0, final_scores=0.6232707,
final_local_scores=0.58662444, pem=0.0, similarity=21.595757]
32: 805862 Hutnik_Nowa_Huta
    [pem_pos=0, final_scores=0.49973565,
final_local_scores=0.5207694, pem=0.0, similarity=19.63051]
33: 8538149 Raków_Częstochowa
    [pem_pos=0, final_scores=0.52505463,
final_local_scores=0.54642546, pem=0.0, similarity=20.39614]
34: 677663 Zagłębie_Lubin
    [pem_pos=0, final_scores=0.5755434,
final_local_scores=0.55021274, pem=0.0, similarity=20.50916]
35: 677685 Ruch_Chorzów
    [pem_pos=0, final_scores=0.61294633,
final_local_scores=0.589275, pem=0.0, similarity=21.674854]
36: 677580 Górnik_Zabrze
    [pem_pos=0, final_scores=0.65329367,
final_local_scores=0.62027705, pem=0.0, similarity=22.600018]
37: 677683 ŁKS_Łódź
    [pem_pos=0, final_scores=0.5104132,
final_local_scores=0.4695084, pem=0.0, similarity=18.10078]
```

Sample: b'343_SOCCER'

SOCCER - [1 1 RUSSIAN 1] PREMIER DIVISION RESULTS /

STANDINGS . [2 2 MOSCOW 2] 1996-08-25 Results of [3 3 Russian 3] premier league matches played on Saturday : Alaniya Vladikavkaz 3 [4 4 Zhemchuzhina Sochi 4] 1 [5 5 Baltika Kalin ingrad 5] 2 [6 6 Zenit St Petersburg 6] 0 Chernomorets Novorossiisk 2 [7 7 Rostselmash Rostov 7] 1 [8 8 Lokomotiv Mo scow 8] 2 [9 9 Torpedo Moscow 9] 1 [10 10 Rotor Volgograd 10] 0 [11 11 Dynamo Moscow 11] 1 [12 12 CSKA Moscow 12] 4 Kamaz Naberezhnye Chelny 2 [13 13 Lada Togliatti 13] 1 [14 14 Spart ak Moscow 14] 1 Tekstilshik Kamyshin 2 [15 15 Krylya Sovetov S amara 15] 1 [16 16 Lokomotiv Nizhny Novgorod 16] 2 [17 17 Ura lmash Yekaterinburg 17] 2 Standings (tabulate under games played , won , drawn , lost , goals for , goals against , points) . Note - if more than one team has the same number of points , precedence is given to the one with most wins . If more than one team has the same number of wins and points , precedence goes to the side with the most successful record against the others) . Alaniya Vladikavkaz 24 16 5 3 48 25 53 [18 18 Dynamo Moscow 18] 25 15 7 3 43 21 52 [19 19 Ro tor Volgograd 19] 23 15 5 3 42 17 50 [20 20 Spartak Moscow 20] 25 14 7 4 48 24 49 [21 21 CSKA Moscow 21] 25 13 6 6 40 27 45 [22 22 Lokomotiv Nizhny Novgorod 22] 25 11 4 10 27 35 37 [23 23 Lokomotiv Moscow 23] 25 9 9 7 30 24 36 [24 24 Baltika Kalinin grad 24] 25 8 10 7 29 26 34 [25 25 Torpedo Moscow 25] 25 8 9 8 31 33 33 [26 26 Zenit St Petersburg 26] 24 9 4 11 24 26 31 [27 27 Krylya Sovetov Samara 27] 25 8 7 10 19 29 31 [28 28 Zhe mchuzhina Sochi 28] 25 8 4 13 26 38 28 [29 29 Rostselmash Ros tov 29] 24 7 7 10 42 38 28 Chernomorets Novorossiisk 25 7 5 13 25 38 26 Kamaz Naberezhnye Chelny 24 5 4 15 25 42 19 [30 30 Lada Togliatti 30] 24 4 6 14 15 37 18 [31 31 Tekstilshchik Ka myshin 31] 25 3 9 13 15 30 18 [32 32 Uralmash Yekaterinburg 32] 24 3 8 13 24 43 16

1: 25391 Russia
2: 19004 Moscow
3: 25391 Russia
4: 5503838 FC_Zhemchuzhina-Sochi
5: 5808153 FC_Baltika_Kaliningrad
6: 891213 FC_Zenit_Saint_Petersburg
7: 1612159 FC_Rostov
8: 1135247 FC_Lokomotiv_Moscow
9: 1443555 FC_Torpedo_Moscow
10: 3108494 FC_Rotor_Volgograd
11: 968856 FC_Dynamo_Moscow
12: 861274 PFC_CSKA_Moscow
13: 5826257 FC_Lada-Togliatti_Togliatti
14: 929190 FC_Spartak_Moscow
15: 1343724 FC_Krylia_Sovetov_Samara
16: 5652475 FC_Lokomotiv_Nizhny_Novgorod
17: 5683235 FC_Ural_Sverdlovsk_Oblast
18: 968856 FC_Dynamo_Moscow
19: 3108494 FC_Rotor_Volgograd
20: 929190 FC_Spartak_Moscow
21: 861274 PFC_CSKA_Moscow

```
22: 5652475 FC_Lokomotiv_Nizhny_Novgorod
23: 1135247 FC_Lokomotiv_Moscow
24: 5808153 FC_Baltika_Kaliningrad
25: 1443555 FC_Torpedo_Moscow
26: 891213 FC_Zenit_Saint_Petersburg
27: 1343724 FC_Krylia_Sovetov_Samara
28: 5503838 FC_Zhemchuzhina-Sochi
29: 1612159 FC_Rostov
30: 5826257 FC_Lada-Togliatti_Togliatti
31: 5702383 FC_Tekstilshchik_Kamyshin
32: 5683235 FC_Ural_Sverdlovsk_Oblast
1: 25391 Russia
    [pem_pos=1, final_scores=0.68978727,
final_local_scores=0.7681447, pem=-1.0515022,
similarity=32.730915]
2: 19004 Moscow
    [pem_pos=0, final_scores=0.6422303,
final_local_scores=0.7058234, pem=-0.106582455,
similarity=25.732504]
3: 25391 Russia
    [pem_pos=1, final_scores=0.63043594,
final_local_scores=0.7033124, pem=-1.0515022,
similarity=30.796188]
4: 5503838 FC_Zhemchuzhina-Sochi
    [pem_pos=0, final_scores=0.5461318,
final_local_scores=0.55717194, pem=0.0, similarity=20.716835]
5: 5808153 FC_Baltika_Kaliningrad
    [pem_pos=0, final_scores=0.45626134,
final_local_scores=0.43297738, pem=0.0, similarity=17.010622]
6: 891213 FC_Zenit_Saint_Petersburg
    [pem_pos=0, final_scores=0.5558216,
final_local_scores=0.5166123, pem=0.0, similarity=19.506454]
7: 1612159 FC_Rostov
    [pem_pos=0, final_scores=0.50021523,
final_local_scores=0.46820652, pem=0.0, similarity=18.06193]
8: 1135247 FC_Lokomotiv_Moscow
    [pem_pos=0, final_scores=0.5505338,
final_local_scores=0.51539564, pem=0.0, similarity=19.470146]
9: 1443555 FC_Torpedo_Moscow
    [pem_pos=0, final_scores=0.739658,
final_local_scores=0.7170592, pem=0.0, similarity=25.48819]
10: 3108494 FC_Rotor_Volgograd
    [pem_pos=0, final_scores=0.44250023,
final_local_scores=0.40701163, pem=0.0, similarity=16.235752]
11: 968856 FC_Dynamo_Moscow
    [pem_pos=0, final_scores=0.5643591,
final_local_scores=0.5175375, pem=-0.45276654,
similarity=21.996279]
12: 861274 PFC_CSKA_Moscow
    [pem_pos=0, final_scores=0.4046021,
final_local_scores=0.35523218, pem=-0.3490741,
similarity=16.58887]
```

13: 5826257 FC_Lada-Togliatti_Togliatti
[pem_pos=1, final_scores=0.4263777,
final_local_scores=0.41470253, pem=-0.292388,
similarity=18.055315]
14: 929190 FC_Spartak_Moscow
[pem_pos=0, final_scores=0.5555888,
final_local_scores=0.5155185, pem=0.0, similarity=19.473812]
15: 1343724 FC_Krylia_Sovetov_Samara
[pem_pos=0, final_scores=0.5283078,
final_local_scores=0.49069792, pem=0.0, similarity=18.733118]
16: 5652475 FC_Lokomotiv_Nizhny_Novgorod
[pem_pos=0, final_scores=0.39518505,
final_local_scores=0.3860793, pem=0.0, similarity=15.61109]
17: 5683235 FC_Ural_Sverdlovsk_Oblast
[pem_pos=0, final_scores=0.57563186,
final_local_scores=0.55233836, pem=0.0, similarity=20.572592]
18: 968856 FC_Dynamo_Moscow
[pem_pos=0, final_scores=0.58100927,
final_local_scores=0.5357252, pem=-0.45276654,
similarity=22.539036]
19: 3108494 FC_Rotor_Volgograd
[pem_pos=0, final_scores=0.47948784,
final_local_scores=0.447415, pem=0.0, similarity=17.44147]
20: 929190 FC_Spartak_Moscow
[pem_pos=0, final_scores=0.5684792,
final_local_scores=0.5295993, pem=0.0, similarity=19.894012]
21: 861274 PFC_CSKA_Moscow
[pem_pos=0, final_scores=0.41051668,
final_local_scores=0.36169297, pem=-0.3490741,
similarity=16.781672]
22: 5652475 FC_Lokomotiv_Nizhny_Novgorod
[pem_pos=0, final_scores=0.3985579,
final_local_scores=0.3897636, pem=0.0, similarity=15.721036]
23: 1135247 FC_Lokomotiv_Moscow
[pem_pos=0, final_scores=0.5360677,
final_local_scores=0.49959368, pem=0.0, similarity=18.998585]
24: 5808153 FC_Baltika_Kaliningrad
[pem_pos=0, final_scores=0.4314115,
final_local_scores=0.4058327, pem=0.0, similarity=16.200571]
25: 1443555 FC_Torpedo_Moscow
[pem_pos=0, final_scores=0.718323,
final_local_scores=0.69375396, pem=0.0, similarity=24.792715]
26: 891213 FC_Zenit_Saint_Petersburg
[pem_pos=0, final_scores=0.522385,
final_local_scores=0.48008794, pem=0.0, similarity=18.416494]
27: 1343724 FC_Krylia_Sovetov_Samara
[pem_pos=0, final_scores=0.52653915,
final_local_scores=0.4887659, pem=0.0, similarity=18.675463]
28: 5503838 FC_Zhemchuzhina-Sochi
[pem_pos=0, final_scores=0.52466613,
final_local_scores=0.533724, pem=0.0, similarity=20.017101]
29: 1612159 FC_Rostov

```

[pem_pos=0, final_scores=0.4448176,
final_local_scores=0.40769303, pem=0.0, similarity=16.256086]
30: 5826257 FC_Lada-Togliatti_Togliatti
    [pem_pos=1, final_scores=0.39174032,
final_local_scores=0.3768664, pem=-0.292388,
similarity=16.926208]
31: 5702383 FC_Tekstilshchik_Kamyshin
    [pem_pos=0, final_scores=0.43621665,
final_local_scores=0.44923097, pem=0.0, similarity=17.49566]
32: 5683235 FC_Ural_Sverdlovsk_Oblast
    [pem_pos=0, final_scores=0.5253447,
final_local_scores=0.49740732, pem=0.0, similarity=18.93334]

```

Sample: b'344_AUSTRALIAN'

AUSTRALIAN RULES-AFL RESULTS AND STANDINGS . [1 1 MELBOURNE 1] 1996-08-25 Results of [2 2 Australian Rules 2] matches played at the weekend . Played Sunday : [3 3 Adelaide 3] 14.12 (96) [4 4 Collingwood 4] 24 . 9 (153) [5 5 West Coast 5] 24 . 7 (151) [6 6 Melbourne 6] 11.12 (78) [7 7 Richmond 7] 28.19 (187) [8 8 Fitzroy 8] 5 . 6 (36) Played Saturday : [9 9 Carlton 9] 13.18 (96) [10 10 Footscray 10] 9.12 (66) [11 11 Essendon 11] 14.16 (100) [12 12 Sydney 12] 12.10 (82) [13 13 St Kilda 13] 9 . 9 (63) [14 14 Hawthorn 14] 12 . 8 (80) [15 15 Brisbane 15] 10.11 (71) [16 16 Fremantle 16] 10.10 (70) Played Friday : [17 17 North Melbourne 17] 14.12 (96) [18 18 Geelong 18] 16.13 (109) Standings (tabulate under played , won , drawn , lost , points for , against , percentage , total points) : [19 19 Brisbane 19] 21 15 1 5 2123 1631 130.2 62 [20 20 Sydney 20] 21 15 1 5 2067 1687 122.5 62 [21 21 West Coast 21] 21 15 0 6 2151 1673 128.6 60 [22 22 North Melbourne 22] 21 15 0 6 2385 1873 127.3 60 [23 23 Carlton 23] 21 14 0 7 2009 1844 108.9 56 [24 24 Geelong 24] 21 13 1 7 2288 1940 117.9 54 [25 25 Essendon 25] 21 13 1 7 2130 1947 109.4 54 [26 26 Richmond 26] 21 11 0 10 2173 1803 120.5 44 [27 27 Hawthorn 27] 21 10 1 10 1791 1820 98.4 42 [28 28 St Kilda 28] 21 9 0 12 1909 1958 97.5 36 [29 29 Collingwood 29] 21 8 0 13 2103 2091 100.6 32 [30 30 Adelaide 30] 21 8 0 13 2158 2183 98.9 32 [31 31 Melbourne 31] 21 7 0 14 1642 2361 69.5 28 [32 32 Fremantle 32] 21 6 0 15 1673 1912 87.5 24 [33 33 Footscray 33] 21 5 1 15 1578 2060 76.6 22 [34 34 Fitzroy 34] 21 1 0 20 1381 2778 49.7 4

- 1: 17306237 Melbourne
- 2: 2403 Australian_rules_football
- 3: 100978 Adelaide_Football_Club
- 4: 90407 Collingwood_Football_Club
- 5: 295230 West_Coast_Eagles
- 6: 97831 Melbourne_Football_Club
- 7: 100040 Richmond_Football_Club
- 8: 210973 Fitzroy_Football_Club
- 9: 89853 Carlton_Football_Club
- 10: 34094 Western_Bulldogs
- 11: 10257 Essendon_Football_Club

```
12: 29210 Sydney_Swans
13: 92922 St_Kilda_Football_Club
14: 12941233 Hawthorn_Football_Club
15: 5026 Brisbane_Lions
16: 312661 Fremantle_Football_Club
17: 22106 North_Melbourne_Football_Club
18: 13007 Geelong_Football_Club
19: 5026 Brisbane_Lions
20: 29210 Sydney_Swans
21: 295230 West_Coast_Eagles
22: 22106 North_Melbourne_Football_Club
23: 89853 Carlton_Football_Club
24: 13007 Geelong_Football_Club
25: 10257 Essendon_Football_Club
26: 100040 Richmond_Football_Club
27: 12941233 Hawthorn_Football_Club
28: 92922 St_Kilda_Football_Club
29: 90407 Collingwood_Football_Club
30: 100978 Adelaide_Football_Club
31: 97831 Melbourne_Football_Club
32: 312661 Fremantle_Football_Club
33: 34094 Western_Bulldogs
34: 210973 Fitzroy_Football_Club
1: 17306237 Melbourne
    [pem_pos=0, final_scores=0.48868144,
final_local_scores=0.5441043, pem=-0.16841523,
similarity=21.242739]
2: 2403 Australian_rules_football
    [pem_pos=0, final_scores=0.4914902,
final_local_scores=0.480344, pem=0.0, similarity=18.424135]
3: 100978 Adelaide_Football_Club
    [pem_pos=1, final_scores=0.60765725,
final_local_scores=0.56256765, pem=-2.7790864,
similarity=35.990955]
4: 90407 Collingwood_Football_Club
    [pem_pos=0, final_scores=0.60913885,
final_local_scores=0.55161744, pem=-0.42581904,
similarity=22.866747]
5: 295230 West_Coast_Eagles
    [pem_pos=2, final_scores=0.6574449,
final_local_scores=0.61705303, pem=-1.9809854,
similarity=33.276714]
6: 97831 Melbourne_Football_Club
    [pem_pos=1, final_scores=0.6899207,
final_local_scores=0.6393284, pem=-2.786373,
similarity=38.321278]
7: 100040 Richmond_Football_Club
    [pem_pos=2, final_scores=0.84510154,
final_local_scores=0.81033695, pem=-2.6048384,
similarity=42.437298]
8: 210973 Fitzroy_Football_Club
    [pem_pos=0, final_scores=0.95674175,
```

```
final_local_scores=0.94421154, pem=-0.51888555,
similarity=35.08865]
9: 89853 Carlton_Football_Club
    [pem_pos=0, final_scores=0.8497898,
final_local_scores=0.83616215, pem=-0.727443,
similarity=32.998405]
10: 34094 Western_Bulldogs
    [pem_pos=0, final_scores=0.6428849,
final_local_scores=0.59227824, pem=-0.42127568,
similarity=24.05544]
11: 10257 Essendon_Football_Club
    [pem_pos=0, final_scores=0.5903477,
final_local_scores=0.5318464, pem=-0.09006156,
similarity=20.45084]
12: 29210 Sydney_Swans
    [pem_pos=1, final_scores=0.57332104,
final_local_scores=0.5142738, pem=-4.0017147,
similarity=41.198616]
13: 92922 St_Kilda_Football_Club
    [pem_pos=0, final_scores=0.75582606,
final_local_scores=0.7138069, pem=-0.48976773,
similarity=28.054567]
14: 12941233 Hawthorn_Football_Club
    [pem_pos=0, final_scores=0.7319385,
final_local_scores=0.68793637, pem=-0.5078096,
similarity=27.380653]
15: 5026 Brisbane_Lions
    [pem_pos=2, final_scores=0.33711565,
final_local_scores=0.27247924, pem=-3.4494781,
similarity=30.979836]
16: 312661 Fremantle_Football_Club
    [pem_pos=1, final_scores=0.6237454,
final_local_scores=0.5809157, pem=-1.2687272,
similarity=28.324932]
17: 22106 North_Melbourne_Football_Club
    [pem_pos=0, final_scores=0.7050305,
final_local_scores=0.65692014, pem=0.0, similarity=23.69352]
18: 13007 Geelong_Football_Club
    [pem_pos=1, final_scores=0.47281963,
final_local_scores=0.40555876, pem=-0.8939384,
similarity=21.053772]
19: 5026 Brisbane_Lions
    [pem_pos=2, final_scores=0.33865815,
final_local_scores=0.27416414, pem=-3.4494781,
similarity=31.030115]
20: 29210 Sydney_Swans
    [pem_pos=1, final_scores=0.6323475,
final_local_scores=0.57875127, pem=-4.0017147,
similarity=43.12275]
21: 295230 West_Coast_Eagles
    [pem_pos=2, final_scores=0.5926263,
final_local_scores=0.54624856, pem=-1.9809854,
```

```
similarity=31.163765]
22: 22106 North_Melbourne_Football_Club
    [pem_pos=0, final_scores=0.69242316,
final_local_scores=0.64314854, pem=0.0, similarity=23.282549]
23: 89853 Carlton_Football_Club
    [pem_pos=0, final_scores=0.7767717,
final_local_scores=0.756401, pem=-0.727443,
similarity=30.618176]
24: 13007 Geelong_Football_Club
    [pem_pos=1, final_scores=0.4179396,
final_local_scores=0.34561068, pem=-0.8939384,
similarity=19.264801]
25: 10257 Essendon_Football_Club
    [pem_pos=0, final_scores=0.5295108,
final_local_scores=0.46539134, pem=-0.09006156,
similarity=18.467688]
26: 100040 Richmond_Football_Club
    [pem_pos=2, final_scores=0.7188011,
final_local_scores=0.6723728, pem=-2.6048384,
similarity=38.320175]
27: 12941233 Hawthorn_Football_Club
    [pem_pos=0, final_scores=0.7558822,
final_local_scores=0.71409124, pem=-0.5078096,
similarity=28.161167]
28: 92922 St_Kilda_Football_Club
    [pem_pos=0, final_scores=0.6302792,
final_local_scores=0.57666606, pem=-0.48976773,
similarity=23.962011]
29: 90407 Collingwood_Football_Club
    [pem_pos=0, final_scores=0.57458717,
final_local_scores=0.51387495, pem=-0.42581904,
similarity=21.740437]
30: 100978 Adelaide_Football_Club
    [pem_pos=1, final_scores=0.570278,
final_local_scores=0.52173644, pem=-2.7790864,
similarity=34.772472]
31: 97831 Melbourne_Football_Club
    [pem_pos=1, final_scores=0.5911874,
final_local_scores=0.5314773, pem=-2.786373,
similarity=35.102783]
32: 312661 Fremantle_Football_Club
    [pem_pos=1, final_scores=0.51107097,
final_local_scores=0.45783603, pem=-1.2687272,
similarity=24.651989]
33: 34094 Western_Bulldogs
    [pem_pos=0, final_scores=0.66943854,
final_local_scores=0.6212841, pem=-0.42127568,
similarity=24.921034]
34: 210973 Fitzroy_Football_Club
    [pem_pos=0, final_scores=0.7784545,
final_local_scores=0.74945974, pem=-0.51888555,
similarity=29.276867]
```

Sample: b'345_RUGBY'

[1 1 RUGBY LEAGUE 1] - [2 2 AUSTRALIAN 2] RUGBY LEAGUE
RESULTS / STANDINGS . [3 3 SYDNEY 3] 1996-08-25 Results of [4
4 Australian 4] rugby league matches played at the weekend .
Played Sunday : [5 5 Sydney Bulldogs 5] 17 [6 6 South Queensl
and 6] 16 [7 7 Brisbane 7] 38 [1 1 Gold Coast 1] 10 [8 8 North
Sydney 8] 46 [9 9 South Sydney 9] 4 [10 10 Illawarra 10] 42 [1
1 11 Penrith 11] 2 [12 12 St George 12] 20 [13 13 North Queen
sland 13] 24 [14 14 Manly 14] 42 [15 15 Western Suburbs 15]
12 Played Saturday : [16 16 Parramatta 16] 14 [17 17 Sydney Ti
gers 17] 26 [18 18 Newcastle 18] 24 [19 19 Western Reds 19]
20 Played Friday : [20 20 Canberra 20] 30 [21 21 Auckland 21]
6 Premiership standings (tabulate under played , won ,
drawn , lost , points for , against , total points) : [22 22
Manly 22] 21 17 0 4 501 181 34 [23 23 Brisbane 23] 21 16 0 5
569 257 32 [24 24 North Sydney 24] 21 14 2 5 560 317 30 [25 25
Sydney City 25] 20 14 1 5 487 293 29 [26 26 Cronulla 26] 20
12 2 6 359 258 26 [27 27 Canberra 27] 21 12 1 8 502 374 25 [28
28 St George 28] 21 12 1 8 421 344 25 [29 29 Newcastle 29] 21
11 1 9 416 366 23 [30 30 Western Suburbs 30] 21 11 1 9 382
426 23 [31 31 Auckland 31] 21 11 0 10 406 389 22 [32 32 Sydney
Tigers 32] 21 11 0 10 309 435 22 [33 33 Parramatta 33] 21 10
1 10 388 391 21 [34 34 Sydney Bulldogs 34] 21 10 0 11 325 356
20 [35 35 Illawarra 35] 21 8 0 13 395 432 16 [36 36 Western Re
ds 36] 21 6 1 14 297 398 13 [37 37 Penrith 37] 21 6 1 14 339
448 13 [38 38 North Queensland 38] 21 6 0 15 266 593 12 [39 39
Gold Coast 39] 21 5 1 15 351 483 11 [40 40 South Sydney 40]
21 5 1 15 304 586 11 [41 41 South Queensland 41] 21 4 0 17
210 460 8

1: 25735 Rugby_league
2: 4689264 Australia
3: 27862 Sydney
4: 4689264 Australia
5: 516739 Canterbury-Bankstown_Bulldogs
6: 2444787 South_Queensland_Crushers
7: 5025 Brisbane_Broncos
8: 693234 North_Sydney_Bears
9: 1079930 South_Sydney_Rabbitohs
10: 853881 Illawarra_Steelers
11: 471314 Penrith_Panthers
12: 853089 St._George_Dragons
13: 995013 North_Queensland_Cowboys
14: 885301 Manly-Warringah_Sea_Eagles
15: 1532351 Western_Suburbs_Magpies
16: 869062 Parramatta_Eels
17: 1464504 Balmain_Tigers
18: 797018 Newcastle_Knights
19: 2015377 WA_Reds
20: 454452 Canberra_Raiders
21: 702650 New_Zealand_Warriors
22: 885301 Manly-Warringah_Sea_Eagles

```
23: 5025 Brisbane_Broncos
24: 693234 North_Sydney_Bears
25: 424878 Sydney_Roosters
26: 525034 Cronulla-Sutherland_Sharks
27: 454452 Canberra_Raiders
28: 853089 St._George_Dragons
29: 797018 Newcastle_Knights
30: 1532351 Western_Suburbs_Magpies
31: 702650 New_Zealand_Warriors
32: 1464504 Balmain_Tigers
33: 869062 Parramatta_Eels
34: 516739 Canterbury-Bankstown_Bulldogs
35: 853881 Illawarra_Steelers
36: 2015377 WA_Reds
37: 471314 Penrith_Panthers
38: 995013 North_Queensland_Cowboys
39: 1613020 Gold_Coast_Titans
40: 1079930 South_Sydney_Rabbitohs
41: 2444787 South_Queensland_Crushers
1: 2444560 Gold_Coast_Chargers
    [winner: 1613020 Gold_Coast_Titans, pem_pos=15,
final_scores=0.2765823, final_local_scores=0.20897424,
pem=-3.7605772, similarity=30.776524]
    gt: 2444560 Gold_Coast_Chargers, pem_pos=13,
final_scores=0.07469396, final_local_scores=0.008825302,
pem=-3.670241, similarity=24.31242 ]
1: 25735 Rugby_league
    [pem_pos=0, final_scores=0.2623,
final_local_scores=0.23075414, pem=0.0, similarity=10.975877]
2: 4689264 Australia
    [pem_pos=0, final_scores=0.6214021,
final_local_scores=0.717624, pem=-0.491009,
similarity=28.175228]
3: 27862 Sydney
    [pem_pos=0, final_scores=0.55993515,
final_local_scores=0.6299165, pem=-0.1443489,
similarity=23.67267]
4: 4689264 Australia
    [pem_pos=0, final_scores=0.60867256,
final_local_scores=0.70371896, pem=-0.491009,
similarity=27.760273]
5: 516739 Canterbury-Bankstown_Bulldogs
    [pem_pos=0, final_scores=0.67023665,
final_local_scores=0.6240777, pem=0.0, similarity=22.713436]
6: 2444787 South_Queensland_Crushers
    [pem_pos=0, final_scores=0.5810339,
final_local_scores=0.58359116, pem=0.0, similarity=21.505238]
7: 5025 Brisbane_Broncos
    [pem_pos=5, final_scores=0.46469963,
final_local_scores=0.40184748, pem=-3.5955102,
similarity=35.634586]
8: 693234 North_Sydney_Bears
```

```
[pem_pos=1, final_scores=0.5321894,
final_local_scores=0.4855256, pem=-1.09192,
similarity=24.516796]
9: 1079930 South_Sydney_Rabbitohs
    [pem_pos=0, final_scores=0.86016124,
final_local_scores=0.8320313, pem=0.0, similarity=28.919188]
10: 853881 Illawarra_Steelers
    [pem_pos=3, final_scores=0.524071,
final_local_scores=0.49182552, pem=-1.9148196,
similarity=29.179853]
11: 471314 Penrith_Panthers
    [pem_pos=2, final_scores=0.61598665,
final_local_scores=0.5669476, pem=-1.61939,
similarity=29.815054]
12: 853089 St._George_Dragons
    [pem_pos=2, final_scores=0.43216664,
final_local_scores=0.37537855, pem=-2.12389,
similarity=26.8418]
13: 995013 North_Queensland_Cowboys
    [pem_pos=1, final_scores=0.6299718,
final_local_scores=0.58749825, pem=-1.0438483,
similarity=27.298443]
14: 885301 Manly-Warringah_Sea_Eagles
    [pem_pos=1, final_scores=0.5308759,
final_local_scores=0.47287613, pem=-1.61336,
similarity=26.974983]
15: 1532351 Western_Suburbs_Magpies
    [pem_pos=0, final_scores=0.39365274,
final_local_scores=0.3322476, pem=-0.31132302,
similarity=15.697667]
16: 869062 Parramatta_Eels
    [pem_pos=1, final_scores=0.7696478,
final_local_scores=0.7355254, pem=-1.3704162,
similarity=33.491795]
17: 1464504 Balmain_Tigers
    [pem_pos=0, final_scores=0.53876597,
final_local_scores=0.49699807, pem=0.0, similarity=18.921127]
18: 797018 Newcastle_Knights
    [pem_pos=14, final_scores=0.5497462,
final_local_scores=0.4966849, pem=-3.8539813,
similarity=39.87033]
19: 2015377 WA_Reds
    [pem_pos=0, final_scores=0.45028335,
final_local_scores=0.43280828, pem=0.0, similarity=17.005575]
20: 454452 Canberra_Raiders
    [pem_pos=1, final_scores=0.74689645,
final_local_scores=0.71455204, pem=-2.8732612,
similarity=41.038612]
21: 702650 New_Zealand_Warriors
    [pem_pos=13, final_scores=0.5836856,
final_local_scores=0.54878485, pem=-3.4653459,
similarity=39.311638]
```

```
22: 885301 Manly-Warringah_Sea_Eagles
    [pem_pos=1, final_scores=0.5387307,
final_local_scores=0.48145622, pem=-1.61336,
similarity=27.231031]
23: 5025 Brisbane_Broncos
    [pem_pos=5, final_scores=0.35763395,
final_local_scores=0.28489447, pem=-3.5955102,
similarity=32.144474]
24: 693234 North_Sydney_Bears
    [pem_pos=1, final_scores=0.5243643,
final_local_scores=0.47697788, pem=-1.09192,
similarity=24.261715]
25: 424878 Sydney_Roosters
    [pem_pos=1, final_scores=0.47885233,
final_local_scores=0.41413927, pem=-0.8080674,
similarity=20.84285]
26: 525034 Cronulla-Sutherland_Sharks
    [pem_pos=1, final_scores=0.58151346,
final_local_scores=0.5314472, pem=-0.7072372,
similarity=23.795221]
27: 454452 Canberra_Raiders
    [pem_pos=1, final_scores=0.721031,
final_local_scores=0.6862979, pem=-2.8732612,
similarity=40.195454]
28: 853089 St_George_Dragons
    [pem_pos=2, final_scores=0.3759752,
final_local_scores=0.31399795, pem=-2.12389,
similarity=25.010082]
29: 797018 Newcastle_Knights
    [pem_pos=14, final_scores=0.51775926,
final_local_scores=0.46174407, pem=-3.8539813,
similarity=38.827625]
30: 1532351 Western_Suburbs_Magpies
    [pem_pos=0, final_scores=0.40997547,
final_local_scores=0.3500777, pem=-0.31132302,
similarity=16.229752]
31: 702650 New_Zealand_Warriors
    [pem_pos=13, final_scores=0.531167,
final_local_scores=0.49141622, pem=-3.4653459,
similarity=37.599644]
32: 1464504 Balmain_Tigers
    [pem_pos=0, final_scores=0.5123736,
final_local_scores=0.46816844, pem=0.0, similarity=18.060793]
33: 869062 Parramatta_Eels
    [pem_pos=1, final_scores=0.78566617,
final_local_scores=0.7530231, pem=-1.3704162,
similarity=34.01396]
34: 516739 Canterbury-Bankstown_Bulldogs
    [pem_pos=0, final_scores=0.6969459,
final_local_scores=0.6532535, pem=0.0, similarity=23.5841]
35: 853881 Illawarra_Steelers
    [pem_pos=3, final_scores=0.49122638,
```

```
final_local_scores=0.45594782, pem=-1.9148196,
similarity=28.109192]
36: 2015377 WA_Reds
    [pem_pos=0, final_scores=0.49811321,
final_local_scores=0.4850551, pem=0.0, similarity=18.564724]
37: 471314 Penrith_Panthers
    [pem_pos=2, final_scores=0.6343643,
final_local_scores=0.5870224, pem=-1.61939,
similarity=30.414127]
38: 995013 North_Queensland_Cowboys
    [pem_pos=1, final_scores=0.6248694,
final_local_scores=0.5819246, pem=-1.0438483,
similarity=27.132113]
39: 1613020 Gold_Coast_Titans
    [pem_pos=15, final_scores=0.35885823,
final_local_scores=0.29883242, pem=-3.7605772,
similarity=33.458073]
40: 1079930 South_Sydney_Rabbitohs
    [pem_pos=0, final_scores=0.889602,
final_local_scores=0.86419076, pem=0.0, similarity=29.87889]
41: 2444787 South_Queensland_Crushers
    [pem_pos=0, final_scores=0.5297341,
final_local_scores=0.5275539, pem=0.0, similarity=19.832973]
1: 1613020 Gold_Coast_Titans
    [pem_pos=15, final_scores=0.2765823,
final_local_scores=0.20897424, pem=-3.7605772,
similarity=30.776524]
```

Sample: b'346_BADMINTON'

BADMINTON - [1 1 MALAYSIAN OPEN 1] RESULTS . [2 2 KUALA LUMPUR 2] 1996-08-25 Results of finals in the [3 3 Malaysian Open 3] badminton tournament on Sunday (prefix numbers denote seedings) : Men 's singles 2 - [4 4 Ong Ewe Hock 4] ([5 5 Malaysia 5]) beat Indra Wijaya ([6 6 Indonesia 6]) 1-15 15-1
15-7 Women 's singles 2 - [7 7 Zhang Ning 7] ([8 8 China 8]) beat 1 - [9 9 Wang Chen 9] ([10 10 China 10]) 11-7
11-8 Women 's doubles 3/4 - [11 11 Marlene Thomsen 11] / [12 12 Lisbet Stuer-Lauridsen 12] ([13 13 Denmark 13]) beat 3/4 - Qiang Hong / Liu Lu ([14 14 China 14]) 10-15 17-14 17-16
Men 's doubles 1 - [15 15 Yap Kim Hock 15] / [16 16 Cheah Soon Kit 16] ([17 17 Malaysia 17]) beat [18 18 Lee Wan Wah 18] / Chong Tan Fook ([19 19 Malaysia 19]) 15-5 15-3

1: 10145751 Malaysia_Open_(badminton)
2: 16854 Kuala_Lumpur
3: 10145751 Malaysia_Open_(badminton)
4: 8278861 Ong_Ewe_Hock
5: 3607937 Malaysia
6: 14579 Indonesia
7: 927560 Zhang_Ning
8: 5405 China
9: 927529 Wang_Chen_(badminton)
10: 5405 China

```
11: 11545746 Marlene_Thomsen
12: 11545810 Lisbeth_Stuer-Lauridsen
13: 76972 Denmark
14: 5405 China
15: 14737607 Yap_Kim_Hock
16: 12250787 Cheah_Soon_Kit
17: 3607937 Malaysia
18: 928104 Lee_Wan_Wah
19: 3607937 Malaysia
1: 10145751 Malaysia_Open_(badminton)
    [pem_pos=2, final_scores=0.38201863,
final_local_scores=0.403759, pem=-1.3940765,
similarity=23.71989]
2: 16854 Kuala_Lumpur
    [pem_pos=0, final_scores=0.44914323,
final_local_scores=0.43568087, pem=0.0, similarity=17.0913]
3: 10145751 Malaysia_Open_(badminton)
    [pem_pos=2, final_scores=0.45139658,
final_local_scores=0.47954386, pem=-1.3940765,
similarity=25.981462]
4: 8278861 Ong_Ewe_Hock
    [pem_pos=0, final_scores=0.44897258,
final_local_scores=0.60311085, pem=0.0, similarity=22.087744]
5: 3607937 Malaysia
    [pem_pos=0, final_scores=0.56854147,
final_local_scores=0.5583328, pem=-0.20223404,
similarity=21.851257]
6: 14579 Indonesia
    [pem_pos=0, final_scores=0.58245856,
final_local_scores=0.5935511, pem=-0.23151328,
similarity=23.061468]
7: 927560 Zhang_Ning
    [pem_pos=0, final_scores=0.65641314,
final_local_scores=0.7110085, pem=0.0, similarity=25.307625]
8: 5405 China
    [pem_pos=0, final_scores=0.45535707,
final_local_scores=0.45874363, pem=-0.15669827,
similarity=18.631687]
9: 927529 Wang_Chen_(badminton)
    [pem_pos=0, final_scores=0.4883728,
final_local_scores=0.5668871, pem=-0.19482312,
similarity=22.066233]
10: 5405 China
    [pem_pos=0, final_scores=0.4239844,
final_local_scores=0.42447376, pem=-0.15669827,
similarity=17.609007]
11: 11545746 Marlene_Thomsen
    [pem_pos=0, final_scores=0.5232432,
final_local_scores=0.6318887, pem=0.0, similarity=22.946531]
12: 11545810 Lisbeth_Stuer-Lauridsen
    [pem_pos=0, final_scores=0.47072285,
final_local_scores=0.5891507, pem=0.0, similarity=21.671146]
```

```
13: 76972 Denmark
    [pem_pos=0, final_scores=0.5393423,
final_local_scores=0.6048585, pem=-0.26762357,
similarity=23.595276]
14: 5405 China
    [pem_pos=0, final_scores=0.4906401,
final_local_scores=0.49728495, pem=-0.15669827,
similarity=19.781837]
15: 14737607 Yap_Kim_Hock
    [pem_pos=0, final_scores=0.47085696,
final_local_scores=0.556941, pem=0.0, similarity=20.709942]
16: 12250787 Cheah_Soon_Kit
    [pem_pos=0, final_scores=0.5936733,
final_local_scores=0.68166095, pem=0.0, similarity=24.431835]
17: 3607937 Malaysia
    [pem_pos=0, final_scores=0.676787,
final_local_scores=0.67657465, pem=-0.20223404,
similarity=25.37983]
18: 928104 Lee_Wan_Wah
    [pem_pos=0, final_scores=0.49999923,
final_local_scores=0.5757891, pem=0.0, similarity=21.272408]
19: 3607937 Malaysia
    [pem_pos=0, final_scores=0.69498414,
final_local_scores=0.69645226, pem=-0.20223404,
similarity=25.973017]
```

Sample: b'347_BASEBALL'

BASEBALL - RESULTS OF [1 1 S. KOREAN 1] PRO-BASEBALL GAMES .
[2 2 SEOUL 2] 1996-08-25 Results of [3 3 South Korean 3] pro-
baseball games played on Saturday . Haitai 10 [4 4 Hanwha 4]
4 Hyundai 5 Samsung 4 Ssangbangwool 4 LG 1 OB 1 Lotte 1 Lotte
1 OB 0* *Note - OB and Lotte played two games . Standings
after games played on Saturday (won , drawn , lost , winning
percentage , games behind first place) W D L PCT GB Haitai
63 2 40 .610 - Ssangbangwool 57 2 47 .547 6 1/2 Hyundai 55 5
47 .537 7 1/2 [5 5 Hanwha 5] 55 1 48 .534 8 Samsung 47 5 54 .
467 15 Lotte 44 6 52 .461 15 1/2 LG 44 5 57 .439 18 OB 40 6
60 .406 21 1/2

```
1: 27019 South_Korea
2: 19159283 Seoul
3: 27019 South_Korea
4: 4086558 Hanwha_Eagles
5: 4086558 Hanwha_Eagles
1: 27019 South_Korea
    [pem_pos=0, final_scores=0.5569053,
final_local_scores=0.5389972, pem=0.0, similarity=20.174463]
2: 19159283 Seoul
    [pem_pos=0, final_scores=0.49760443,
final_local_scores=0.49343157, pem=-0.0378014,
similarity=19.020266]
3: 27019 South_Korea
    [pem_pos=0, final_scores=0.52890307,
```

```
final_local_scores=0.5084089, pem=0.0, similarity=19.261648]
4: 4086558 Hanwha_Eagles
    [pem_pos=1, final_scores=0.5059474,
final_local_scores=0.53260815, pem=-0.6931472,
similarity=23.753244]
5: 4086558 Hanwha_Eagles
    [pem_pos=1, final_scores=0.6783738,
final_local_scores=0.7209578, pem=-0.6931472,
similarity=29.373974]
```

Sample: b'348_BASEBALL'

```
BASEBALL - [1 1 MAJOR LEAGUE 1] RESULTS SUNDAY . [2 2 NEW YO
RK 2] 1996-08-25 Results of [3 3 Major League 3] Baseball
games played on Sunday ( home team in CAPS ) : [4 4 American L
league 4] [5 5 BOSTON 5] 8 [6 6 Seattle 6] 5 [7 7 CLEVELAND 7]
8 [8 8 Milwaukee 8] 5 California 13 [9 9 BALTIMORE 9] 0 [10 10
Oakland 10] 6 [11 11 NEW YORK 11] 4 1 CHICAGO 10 [12 12 Toron
to 12] 9 ( 10 ) [13 13 Texas 13] 13 [14 14 MINNESOTA 14] 2 [15
15 Detroit 15] 7 [16 16 KANSAS CITY 16] 4
1: 38776 Major_League_Baseball
2: 645042 New_York_City
3: 38776 Major_League_Baseball
4: 64777 American_League
5: 4339 Boston_Red_Sox
6: 27660 Seattle_Mariners
7: 6652 Cleveland_Indians
8: 20152 Milwaukee_Brewers
9: 4340 Baltimore_Orioles
10: 22523 Oakland_Athletics
11: 4848143 New_York_Yankees
12: 30847 Toronto_Blue_Jays
13: 30857 Texas_Rangers_(baseball)
14: 20050 Minnesota_Twins
15: 8579 Detroit_Tigers
16: 16883 Kansas_City_Royals
1: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.75952137, final_local_scores=-0.8507015,
pem=-4.591283, similarity=3.671192]
    RECALL MISS: 5945 Chicago_White_Sox]
1: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.5626669,
final_local_scores=0.5095744, pem=-0.47792417,
similarity=21.895454]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.29276413,
final_local_scores=0.36016643, pem=-1.7906687,
similarity=24.575731]
3: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.5703885,
final_local_scores=0.5180091, pem=-0.47792417,
similarity=22.147163]
```

```
4: 64777 American_League
    [pem_pos=0, final_scores=0.5402792,
final_local_scores=0.5097762, pem=0.0, similarity=19.30245]
5: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.052025,
final_local_scores=1.0379465, pem=-4.418931,
similarity=59.094936]
6: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.5292776,
final_local_scores=0.47371787, pem=-3.8567436,
similarity=39.19997]
7: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.1580105,
final_local_scores=1.1483171, pem=-4.191953,
similarity=61.154278]
8: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.3863895,
final_local_scores=0.31261164, pem=-3.40205,
similarity=31.919546]
9: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.1995907,
final_local_scores=1.1945858, pem=-4.046561,
similarity=61.744358]
10: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.2795118,
final_local_scores=1.2832341, pem=-3.9188845,
similarity=63.69548]
11: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.47461385,
final_local_scores=0.40705997, pem=-5.0692368,
similarity=43.804493]
12: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.93791884,
final_local_scores=0.9197076, pem=-4.422637,
similarity=55.586613]
13: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.0383264,
final_local_scores=1.0232697, pem=-4.928425,
similarity=61.42766]
14: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=1.092872,
final_local_scores=1.080399, pem=-4.2906566,
similarity=59.664234]
15: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.1179487,
final_local_scores=1.1060076, pem=-4.1401463,
similarity=59.609947]
16: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.63278216,
final_local_scores=0.5805049, pem=-3.5915844,
similarity=40.944733]
```

Sample: b'349_BASEBALL'

BASEBALL - [1 1 MAJOR LEAGUE 1] STANDINGS AFTER SATURDAY 'S GAMES . [2 2 NEW YORK 2] 1996-08-25 [3 3 Major League Baseball 3] standings after games played on Saturday (tabulate under won , lost , winning percentage and games behind) : [4 4 AMERICAN LEAGUE EASTERN DIVISION 4] W L PCT GB [5 5 NEW YORK 5] 74 54 .578 - [6 6 BALTIMORE 6] 68 60 .531 6 [7 7 BOSTON 7] 65 65 .500 10 [8 8 TORONTO 8] 61 69 .469 14 [9 9 DETROIT 9] 46 83 .357 28 1/2 [10 10 CENTRAL DIVISION 10] [11 11 CLEVELAND 11] 76 53 .589 - 1 CHICAGO 69 62 .527 8 [12 12 MINNESOTA 12] 65 64 .504 11 [13 13 MILWAUKEE 13] 62 68 .477 14 1/2 [14 14 KANSAS CITY 14] 59 72 .450 18 [15 15 WESTERN DIVISION 15] [16 16 TEXAS 16] 74 56 .569 - [17 17 SEATTLE 17] 66 62 .516 7 [18 18 OAKLAND 18] 62 70 .470 13 2 CALIFORNIA 60 69 .465 13 1/2 SUNDAY , AUGUST 25TH SCHEDULE [19 19 SEATTLE 19] AT [20 20 BOSTON 20] [21 21 MILWAUKEE 21] AT [22 22 CLEVELAND 22] 3 CALIFORNIA AT [23 23 BALTIMORE 23] [24 24 OAKLAND 24] AT [1 4 NEW YORK 1] [25 25 TORONTO 25] AT 5 CHICAGO [26 26 TEXAS 26] AT [27 27 MINNESOTA 27] [28 28 DETROIT 28] AT [29 29 KANSAS CITY 29] [30 30 NATIONAL LEAGUE EASTERN DIVISION 30] W L PCT GB [31 31 ATLANTA 31] 81 47 .633 - [32 32 MONTREAL 32] 70 58 .547 11 6 FLORIDA 60 70 .462 22 [33 33 NEW YORK 33] 59 71 .454 23 [34 34 PHILADELPHIA 34] 53 77 .408 29 [35 35 CENTRAL DIVISION 35] [36 36 HOUSTON 36] 69 61 .531 - ST LOUIS 68 61 .527 1/2 [37 37 CINCINNATI 37] 64 64 .500 4 7 CHICAGO 63 64 .496 4 1/2 [38 38 PITTSBURGH 38] 55 74 .426 13 1/2 [39 39 WESTERN DIVISION 39] [40 40 SAN DIEGO 40] 71 60 .542 - 8 LOS ANGELES 69 60 .535 1 [41 41 COLORADO 41] 67 63 .515 3 1/2 [42 42 SAN FRANCISCO 42] 54 73 .425 15 SUNDAY , AUGUST 25TH SCHEDULE 9 CHICAGO AT [43 43 ATLANTA 43] [44 44 PITTSBURGH 44] AT [45 45 COLORADO 45] [46 46 NEW YORK 46] AT 10 LOS ANGELES [47 47 PHILADELPHIA 47] AT [48 48 SAN DIEGO 48] [49 49 MONTREAL 49] AT [50 50 SAN FRANCISCO 50] [51 51 CINCINNATI 51] AT 11 FLORIDA ST LOUIS AT [52 52 HOU STON 52]

1: 38776 Major_League_Baseball
2: 645042 New_York_City
3: 38776 Major_League_Baseball
4: 2004582 American_League_East
5: 4848143 New_York_Yankees
6: 4340 Baltimore_Orioles
7: 4339 Boston_Red_Sox
8: 30847 Toronto_Blue_Jays
9: 8579 Detroit_Tigers
10: 2456377 National_League_Central
11: 6652 Cleveland_Indians
12: 20050 Minnesota_Twins
13: 20152 Milwaukee_Brewers
14: 16883 Kansas_City_Royals
15: 2456474 National_League_West
16: 30857 Texas_Rangers_(baseball)
17: 27660 Seattle_Mariners
18: 22523 Oakland_Athletics

19: 27660 Seattle_Mariners
20: 4339 Boston_Red_Sox
21: 20152 Milwaukee_Brewers
22: 6652 Cleveland_Indians
23: 4340 Baltimore_Orioles
24: 22523 Oakland_Athletics
25: 30847 Toronto_Blue_Jays
26: 30857 Texas_Rangers_(baseball)
27: 20050 Minnesota_Twins
28: 8579 Detroit_Tigers
29: 16883 Kansas_City_Royals
30: 2456294 National_League_East
31: 2140 Atlanta_Braves
32: 20153 Montreal_Expos
33: 4848143 New_York_Yankees
34: 23741 Philadelphia_Phillies
35: 2456377 National_League_Central
36: 13894 Houston_Astros
37: 6671 Cincinnati_Reds
38: 24270 Pittsburgh_Pirates
39: 2456474 National_League_West
40: 28417 San_Diego_Padres
41: 6669 Colorado_Rockies
42: 28416 San_Francisco_Giants
43: 2140 Atlanta_Braves
44: 24270 Pittsburgh_Pirates
45: 6669 Colorado_Rockies
46: 4848143 New_York_Yankees
47: 23741 Philadelphia_Phillies
48: 28417 San_Diego_Padres
49: 20153 Montreal_Expos
50: 28416 San_Francisco_Giants
51: 6671 Cincinnati_Reds
52: 13894 Houston_Astros
1: 5945 Chicago_White_Sox
 [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.7396212, final_local_scores=-0.83030623,
pem=-4.591283, similarity=4.279827]
 RECALL MISS: 5945 Chicago_White_Sox]
2: 1360083 Los_Angeles_Angels_of_Anheim
 [winner: 799305 California_State_Route_60, pem_pos=0,
final_scores=-0.5123114, final_local_scores=-0.43560168,
pem=0.0, similarity=-8.90951]
 RECALL MISS: 1360083 Los_Angeles_Angels_of_Anheim]
3: 1360083 Los_Angeles_Angels_of_Anheim
 [winner: 5407 California, pem_pos=0,
final_scores=-0.8316681, final_local_scores=-0.8659364,
pem=-0.08662026, similarity=-21.2805]
 RECALL MISS: 1360083 Los_Angeles_Angels_of_Anheim]
4: 645042 New_York_City
 [winner: 4848143 New_York_Yankees, pem_pos=11,
final_scores=0.27459377, final_local_scores=0.18651992,

```
pem=-5.0692368, similarity=37.223137 |
    gt: 645042 New_York_City, pem_pos=1,
final_scores=0.05289261, final_local_scores=0.08735287,
pem=-1.7906687, similarity=16.434431 ]
5: 5945 Chicago_White_Sox
    [winner: 6886 Chicago, pem_pos=0,
final_scores=-0.49863625, final_local_scores=-0.54273117,
pem=-0.20658566, similarity=-10.983023|
     RECALL MISS: 5945 Chicago_White_Sox]
6: 11278 Miami_Marlins
    [winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.7794164, final_local_scores=-0.85308355,
pem=-4.620281, similarity=3.757806|
     RECALL MISS: 11278 Miami_Marlins]
7: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.790715, final_local_scores=-0.8861184,
pem=-4.591283, similarity=2.6142807|
     RECALL MISS: 5945 Chicago_White_Sox]
8: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 18213 Los_Angeles_Dodgers, pem_pos=15,
final_scores=-0.12602027, final_local_scores=-0.2530889,
pem=-4.7524934, similarity=22.381834|
     RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
9: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.6979571, final_local_scores=-0.7847945,
pem=-4.591283, similarity=5.637989|
     RECALL MISS: 5945 Chicago_White_Sox]
10: 18110 Los_Angeles
    [winner: 18110 Los_Angeles, pem_pos=0,
final_scores=0.022096023, final_local_scores=0.04519956,
pem=-0.100652285, similarity=5.985921|
     gt: 18110 Los_Angeles, pem_pos=0,
final_scores=0.022096023, final_local_scores=0.04519956,
pem=-0.100652285, similarity=5.985921 ]
11: 11278 Miami_Marlins
    [winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.6472319, final_local_scores=-0.708692,
pem=-4.620281, similarity=8.066737|
     RECALL MISS: 11278 Miami_Marlins]
1: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.59078145,
final_local_scores=0.5432236, pem=-0.47792417,
similarity=22.899614]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.37184596,
final_local_scores=0.435578, pem=-1.7906687,
similarity=26.826164]
3: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.54013485,
final_local_scores=0.48789996, pem=0.0, similarity=18.64962]
```

```
4: 2004582 American_League_East
    [pem_pos=0, final_scores=0.4761005,
final_local_scores=0.4748345, pem=0.0, similarity=18.259722]
5: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.3259676,
final_local_scores=0.24263805, pem=-5.0692368,
similarity=38.897816]
6: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.3166685,
final_local_scores=1.3225996, pem=-4.046561,
similarity=65.56455]
7: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.1675668,
final_local_scores=1.1635084, pem=-4.418931,
similarity=62.841957]
8: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.76417476,
final_local_scores=0.72822285, pem=-4.422637,
similarity=49.872322]
9: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.1797763,
final_local_scores=1.1733799, pem=-4.1401463,
similarity=61.620472]
10: 2456377 National_League_Central
    [pem_pos=0, final_scores=0.5884387,
final_local_scores=0.6083434, pem=-0.98951703,
similarity=27.625042]
11: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.2699597,
final_local_scores=1.2708163, pem=-4.191953,
similarity=64.8099]
12: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=1.2303972,
final_local_scores=1.2309992, pem=-4.2906566,
similarity=64.15844]
13: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.5205032,
final_local_scores=0.4571728, pem=-3.40205,
similarity=36.23354]
14: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.8538745,
final_local_scores=0.82193315, pem=-3.5915844,
similarity=48.149433]
15: 2456474 National_League_West
    [pem_pos=1, final_scores=0.56614107,
final_local_scores=0.5705425, pem=-1.2952092,
similarity=28.159388]
16: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.2446196,
final_local_scores=1.2476494, pem=-4.928425,
similarity=68.123604]
17: 27660 Seattle_Mariners
```

```
[pem_pos=4, final_scores=0.7352671,
final_local_scores=0.69746476, pem=-3.8567436,
similarity=45.87702]
18: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.3004255,
final_local_scores=1.3055911, pem=-3.9188845,
similarity=64.36266]
19: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.68550223,
final_local_scores=0.64310414, pem=-3.8567436,
similarity=44.254795]
20: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.0087012,
final_local_scores=0.98997205, pem=-4.418931,
similarity=57.66329]
21: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.5147934,
final_local_scores=0.45093566, pem=-3.40205,
similarity=36.04741]
22: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.2150083,
final_local_scores=1.2107902, pem=-4.191953,
similarity=63.0186]
23: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.0414304,
final_local_scores=1.0219438, pem=-4.046561,
similarity=56.592384]
24: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.2190206,
final_local_scores=1.2166686, pem=-3.9188845,
similarity=61.709034]
25: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.6536658,
final_local_scores=0.60750866, pem=-4.422637,
similarity=46.26997]
26: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.2359698,
final_local_scores=1.2382009, pem=-4.928425,
similarity=67.84164]
27: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=1.0056005,
final_local_scores=0.98544294, pem=-4.2906566,
similarity=56.830555]
28: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.1945658,
final_local_scores=1.1895351, pem=-4.1401463,
similarity=62.102577]
29: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.77159363,
final_local_scores=0.73205364, pem=-3.5915844,
similarity=45.467247]
30: 2456294 National_League_East
```

```
[pem_pos=0, final_scores=0.4341591,
final_local_scores=0.42828125, pem=0.0, similarity=16.87048]
31: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=1.1743933,
final_local_scores=1.1669106, pem=-4.136474,
similarity=61.407444]
32: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.6812653,
final_local_scores=0.63477135, pem=-4.5098605,
similarity=47.55788]
33: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.54969734,
final_local_scores=0.48688906, pem=-5.0692368,
similarity=46.186752]
34: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.2512875,
final_local_scores=1.2524621, pem=-4.5189333,
similarity=66.040344]
35: 2456377 National_League_Central
    [pem_pos=0, final_scores=0.58105725,
final_local_scores=0.60028034, pem=-0.98951703,
similarity=27.384422]
36: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.3359971,
final_local_scores=1.3451805, pem=-4.650927,
similarity=69.52505]
37: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.0588542,
final_local_scores=1.0421376, pem=-3.771247,
similarity=55.697807]
38: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.2814741,
final_local_scores=1.283498, pem=-4.03702,
similarity=64.345795]
39: 2456474 National_League_West
    [pem_pos=1, final_scores=0.5777002,
final_local_scores=0.58316904, pem=-1.2952092,
similarity=28.536188]
40: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.9257713,
final_local_scores=0.9012049, pem=-3.9307637,
similarity=52.359566]
41: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.2220664,
final_local_scores=1.2280016, pem=-4.482383,
similarity=65.111626]
42: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.98381454,
final_local_scores=0.9590177, pem=-4.409061,
similarity=56.68587]
43: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=0.92534596,
```

```
final_local_scores=0.89486426, pem=-4.136474,
similarity=53.289036]
44: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.2212603,
final_local_scores=1.2177236, pem=-4.03702,
similarity=62.382954]
45: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.2737805,
final_local_scores=1.2844913, pem=-4.482383,
similarity=66.797386]
46: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.58162683,
final_local_scores=0.5217672, pem=-5.0692368,
similarity=47.227585]
47: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.1432544,
final_local_scores=1.1344523, pem=-4.5189333,
similarity=62.518696]
48: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.95328563,
final_local_scores=0.93126017, pem=-3.9307637,
similarity=53.256477]
49: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.6891876,
final_local_scores=0.6434252, pem=-4.5098605,
similarity=47.816128]
50: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.9506555,
final_local_scores=0.9227964, pem=-4.409061,
similarity=55.604958]
51: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.1099839,
final_local_scores=1.0979891, pem=-3.771247,
similarity=57.364525]
52: 13894 Houston_Astros
    [pem_pos=11, final_scores=0.84362704,
final_local_scores=0.80734086, pem=-4.650927,
similarity=53.47484]
1: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.27459377,
final_local_scores=0.18651992, pem=-5.0692368,
similarity=37.223137]
```

Sample: b'350_BASEBALL'

BASEBALL - [1 1 ORIOLES 1] SNEAK PAST [2 2 ANGELS 2] . [3 3 B
ALTIMORE 3] 1996-08-25 [4 4 Rafael Palmeiro 4] 's two-out
single in the sixth inning scored [5 5 Roberto Alomar 5] with
the go-ahead run as the [6 6 Baltimore Orioles 6] rallied
past the [7 7 California Angels 7] 5-4 and took over the [8 8
American League 8] 's wild-card berth on Saturday . The [9 9 0
rioles 9] trailed 4-3 when pinch-hitter [10 10 Mike Devereaux
10] led off the sixth with a triple against reliever [11 11 Ky

le Abbott 11] (0-1) and scored the tying run on [12 12 Alomar 12] 's single . After [13 13 Brady Anderson 13] sacrificed , [14 14 Palmeiro 14] hit the first pitch into right field for a single , scoring [15 15 Alomar 15] . In [16 16 Boston 16] , former Mariner Darren Bragg 's first career grand slam in the sixth inning off reliever [17 17 Randy Johnson 17] lifted the [18 18 Boston Red Sox 18] to their fifth win in six games , a 9-5 victory over [19 19 Seattle 19] . " Just one of those things , I was just trying to make contact , " said [20 20 Bragg 20] . " The bases were loaded and I had two strikes . I was just trying to put the ball in play . I got the good part of the bat on it and it carried out . " In [21 21 Cleveland 21] , [22 22 Kevin Seitzer 22] 's two-out single in the top of the 10th brought home [23 23 David Hulse 23] with the winning run as the [24 24 Milwaukee Brewers 24] sent the [25 25 Cleveland Indians 25] to their third straight extra-inning defeat 4-3 . [26 26 Bob Wickman 26] (5-1) , acquired from the [27 27 New York Yankees 27] on Friday , earned the win in his [28 28 Milwaukee 28] debut despite allowing the tying run in the eighth inning . At [29 29 Minnesota 29] , [30 30 Marty Cordova 30] and [31 31 Matt Lawton 31] hit solo homers and [32 32 Frankie Rodriguez 32] allowed six hits over seven innings to earn his first win as a starter in a month as the [33 33 Minnesota Twins 33] held on to beat the [34 34 Texas Rangers 34] 6-5 . " Yeah , you know it 's fun , it 's always fun when you 've got a chance to go to the ballpark and win a game that 's important , " said Rodriguez . " Every game should be important , but it 's a little more important now . " In [35 35 New York 35] , [36 36 Wally Whitehurst 36] allowed two runs over seven innings for his first win in more than two years and Paul O'Neill 's three-run double snapped a sixth-inning tie as the [37 37 New York Yankees 37] held on for a 5-4 victory over the [38 38 Oakland Athletics 38] . [39 39 Whitehurst 39] , promoted from Triple-A Columbus on Wednesday , allowed seven hits and struck out one without a walk . It was his first win since defeating the [40 40 St. Louis Cardinals 40] on May 28th , 1994 when he was with the [41 41 San Diego Padres 41] . In [42 42 Kansas City 42] , [43 43 Jose Rosado 43] came within one out of his third complete game and [44 44 Michael Tucker 44] homered and drove in three runs as the [45 45 Kansas City Royals 45] broke a six-game losing streak with a 9-2 victory over the [46 46 Detroit Tigers 46] in a battle of cellar-dwellers . [47 47 Rosado 47] (5-3) allowed two runs -- one earned -- and seven hits over 8-2/3 innings with three walks and six strikeouts . In his last four starts , the 21-year-old left-hander has given up only four earned runs in 29-2/3 innings .

1: 4340 Baltimore_Orioles

2: 1360083 Los_Angeles_Angels_of_Anheim

3: 26997138 Baltimore

4: 230185 Rafael_Palmeiro

```
5: 384099 Roberto_Alomar
6: 4340 Baltimore_Orioles
7: 1360083 Los_Angeles_Angels_of_Anheim
8: 64777 American_League
9: 4340 Baltimore_Orioles
10: 4440986 Mike_Devereaux
11: 15475685 Kyle_Abbott
12: 384099 Roberto_Alomar
13: 390266 Brady_Anderson
14: 230185 Rafael_Palmeiro
15: 384099 Roberto_Alomar
16: 24437894 Boston
17: 386384 Randy_Johnson
18: 4339 Boston_Red_Sox
19: 27660 Seattle_Mariners
20: 4233756 Darren_Bragg
21: 5951 Cleveland
22: 1647990 Kevin_Seitzer
23: 18940106 David_Hulse_(baseball)
24: 20152 Milwaukee_Brewers
25: 6652 Cleveland_Indians
26: 1567114 Bob_Wickman
27: 4848143 New_York_Yankees
28: 20152 Milwaukee_Brewers
29: 19590 Minnesota
30: 1547330 Marty_Cordova
31: 1637805 Matt_Lawton
32: 6230681 Frank_Rodriguez
33: 20050 Minnesota_Twins
34: 30857 Texas_Rangers_(baseball)
35: 645042 New_York_City
36: 9580179 Wally_Whitehurst
37: 4848143 New_York_Yankees
38: 22523 Oakland_Athletics
39: 9580179 Wally_Whitehurst
40: 23408844 St._Louis_Cardinals
41: 28417 San_Diego_Padres
42: 17454 Kansas_City,_Missouri
43: 2407601 José_Rosado
44: 1649286 Michael_Tucker_(baseball)
45: 16883 Kansas_City_Royals
46: 8579 Detroit_Tigers
47: 2407601 José_Rosado
1: 4340 Baltimore_Orioles
    [pem_pos=0, final_scores=0.722857,
final_local_scores=0.6806769, pem=-0.18171716,
similarity=25.390675]
2: 1360083 Los_Angeles_Angels_of_Anheim
    [pem_pos=2, final_scores=0.43794078,
final_local_scores=0.37075818, pem=-1.7538311,
similarity=24.691483]
3: 26997138 Baltimore
```

```
[pem_pos=0, final_scores=0.5417176,
final_local_scores=0.60530144, pem=-0.19949137,
similarity=23.23798]
4: 230185 Rafael_Palmeiro
    [pem_pos=0, final_scores=0.47758228,
final_local_scores=0.4877013, pem=0.0, similarity=18.643692]
5: 384099 Roberto_Alomar
    [pem_pos=0, final_scores=0.4703337,
final_local_scores=0.46394825, pem=0.0, similarity=17.934855]
6: 4340 Baltimore_Orioles
    [pem_pos=0, final_scores=0.4542467,
final_local_scores=0.3872609, pem=-0.21811225,
similarity=16.83248]
7: 1360083 Los_Angeles_Angels_of_Anheim
    [pem_pos=0, final_scores=0.51967853,
final_local_scores=0.46004426, pem=-0.016987266,
similarity=17.91073]
8: 64777 American_League
    [pem_pos=0, final_scores=0.47653383,
final_local_scores=0.44389445, pem=0.0, similarity=17.336409]
9: 4340 Baltimore_Orioles
    [pem_pos=0, final_scores=0.52039844,
final_local_scores=0.45952165, pem=-0.18171716,
similarity=18.790962]
10: 4440986 Mike_Devereaux
    [pem_pos=0, final_scores=0.7769025,
final_local_scores=0.83571005, pem=0.0, similarity=29.028967]
11: 15475685 Kyle_Abbott
    [pem_pos=0, final_scores=0.63095254,
final_local_scores=0.6800718, pem=0.0, similarity=24.38441]
12: 384099 Roberto_Alomar
    [pem_pos=0, final_scores=0.2769729,
final_local_scores=0.25273097, pem=0.0, similarity=11.631709]
13: 390266 Brady_Anderson
    [pem_pos=0, final_scores=0.5185502,
final_local_scores=0.53861994, pem=0.0, similarity=20.163206]
14: 230185 Rafael_Palmeiro
    [pem_pos=0, final_scores=0.3269859,
final_local_scores=0.32319757, pem=0.0, similarity=13.734573]
15: 384099 Roberto_Alomar
    [pem_pos=0, final_scores=0.32511622,
final_local_scores=0.30532017, pem=0.0, similarity=13.201076]
16: 24437894 Boston
    [pem_pos=0, final_scores=0.5730652,
final_local_scores=0.6517779, pem=-0.23164093,
similarity=24.799767]
17: 386384 Randy_Johnson
    [pem_pos=0, final_scores=0.5450403,
final_local_scores=0.5358736, pem=0.0, similarity=20.08125]
18: 4339 Boston_Red_Sox
    [pem_pos=0, final_scores=0.50128996,
final_local_scores=0.44383895, pem=-0.08447198,
```

```
similarity=17.794125]
19: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.6027495,
final_local_scores=0.5558503, pem=-3.8567436,
similarity=41.650967]
20: 4233756 Darren_Bragg
    [pem_pos=0, final_scores=1.2493085,
final_local_scores=1.3707676, pem=0.0, similarity=44.996155]
21: 5951 Cleveland
    [pem_pos=0, final_scores=0.5415423,
final_local_scores=0.5942969, pem=-0.57246673,
similarity=24.93788]
22: 1647990 Kevin_Seitzer
    [pem_pos=0, final_scores=0.64609194,
final_local_scores=0.6792189, pem=0.0, similarity=24.35896]
23: 18940106 David_Hulse_(baseball)
    [pem_pos=1, final_scores=0.3833614,
final_local_scores=0.47708756, pem=-0.074108,
similarity=18.729967]
24: 20152 Milwaukee_Brewers
    [pem_pos=0, final_scores=0.43968844,
final_local_scores=0.37322438, pem=-0.12176443,
similarity=15.889647]
25: 6652 Cleveland_Indians
    [pem_pos=0, final_scores=0.5970389,
final_local_scores=0.5413355, pem=-0.08813761,
similarity=20.72355]
26: 1567114 Bob_Wickman
    [pem_pos=0, final_scores=0.68174195,
final_local_scores=0.7240722, pem=0.0, similarity=25.697472]
27: 4848143 New_York_Yankees
    [pem_pos=0, final_scores=0.6178674,
final_local_scores=0.56843793, pem=-0.14222674,
similarity=21.826485]
28: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.5080568,
final_local_scores=0.44790643, pem=-3.40205,
similarity=35.957012]
29: 19590 Minnesota
    [pem_pos=0, final_scores=0.5199731,
final_local_scores=0.6119347, pem=-0.21388707,
similarity=23.514215]
30: 1547330 Marty_Cordova
    [pem_pos=0, final_scores=0.64030623,
final_local_scores=0.695157, pem=0.0, similarity=24.834583]
31: 1637805 Matt_Lawton
    [pem_pos=0, final_scores=0.6149332,
final_local_scores=0.6227419, pem=0.0, similarity=22.673573]
32: 6230681 Frank_Rodriguez
    [pem_pos=0, final_scores=0.5043165,
final_local_scores=0.541505, pem=0.0, similarity=20.249302]
33: 20050 Minnesota_Twins
```

```
[pem_pos=0, final_scores=0.5357902,
final_local_scores=0.480415, pem=-0.052076127,
similarity=18.709452]
34: 30857 Texas_Rangers_(baseball)
[pem_pos=0, final_scores=0.47686166,
final_local_scores=0.41464764, pem=-0.15069729,
similarity=17.283142]
35: 645042 New_York_City
[pem_pos=1, final_scores=0.4384914,
final_local_scores=0.5094269, pem=-1.7906687,
similarity=29.029963]
36: 9580179 Wally_Whitehurst
[pem_pos=0, final_scores=0.7083221,
final_local_scores=0.74061316, pem=0.0, similarity=26.191088]
37: 4848143 New_York_Yankees
[pem_pos=0, final_scores=0.57681006,
final_local_scores=0.523589, pem=-0.14222674,
similarity=20.488104]
38: 22523 Oakland_Athletics
[pem_pos=0, final_scores=0.65366143,
final_local_scores=0.60514843, pem=-0.06842491,
similarity=22.520655]
39: 9580179 Wally_Whitehurst
[pem_pos=0, final_scores=1.278748,
final_local_scores=1.3679576, pem=0.0, similarity=44.912296]
40: 23408844 St._Louis_Cardinals
[pem_pos=0, final_scores=0.45758337,
final_local_scores=0.39622408, pem=-0.2366796,
similarity=17.20093]
41: 28417 San_Diego_Padres
[pem_pos=0, final_scores=0.6040178,
final_local_scores=0.55507267, pem=-0.05076447,
similarity=20.930254]
42: 17454 Kansas_City,_Missouri
[pem_pos=0, final_scores=0.2615049,
final_local_scores=0.27546078, pem=-0.8046767,
similarity=16.685968]
43: 2407601 José_Rosado
[pem_pos=0, final_scores=0.4315936,
final_local_scores=0.50811106, pem=0.0, similarity=19.25276]
44: 1649286 Michael_Tucker_(baseball)
[pem_pos=1, final_scores=0.69741327,
final_local_scores=0.72016865, pem=-0.5160794,
similarity=28.387503]
45: 16883 Kansas_City_Royals
[pem_pos=0, final_scores=0.5082903,
final_local_scores=0.4462369, pem=-0.029036561,
similarity=17.564219]
46: 8579 Detroit_Tigers
[pem_pos=0, final_scores=0.5097934,
final_local_scores=0.45111442, pem=-0.08904677,
similarity=18.036118]
```

47: 2407601 José_Rosado
[pem_pos=0, final_scores=0.5093575,
final_local_scores=0.5996142, pem=0.0, similarity=21.983397]

Sample: b'351_BASEBALL'
BASEBALL - [1 1 BRAVES 1] RALLY TO BEAT [2 2 CUBS 2] . [3 3 A
TLANTA 3] 1996-08-25 [4 4 Fred McGriff 4] went 5-for-5 and
homered twice , including a three-run blast with two out in
the bottom of the ninth inning that lifted the [5 5 Atlanta Br
aves 5] to a 6-5 victory over the [6 6 Chicago Cubs 6] on
Saturday . " I was just trying to hang in there and hit it up
the middle , " said [7 7 McGriff 7] about his homer in the
ninth . " I was just looking for the ball , trying to stay on
it . [8 8 Brad Clontz 8] (6-2) picked up the win in relief
for [9 9 Atlanta 9] , which has won 11 of its last 13 games .
In [10 10 Colorado 10] , [11 11 Mark Thompson 11] threw an
eight-hitter for his third complete game and [12 12 Ellis Bur
ks 12] homered and drove in three runs as the [13 13 Colorado
Rockies 13] beat the [14 14 Pittsburgh Pirates 14] 9-3 . [15 1
5 Vinny Castilla 15] and [16 16 Dante Bichette 16] each added
two RBI for [17 17 Colorado 17] , which improved the major
league 's best home mark to 44-20 . At 1 Florida , [18 18 Kev
in Brown 18] scattered seven hits over eight innings and [19 1
9 Kurt Abbott 19] snapped a sixth-inning tie with a two-run
double as the [20 20 Florida Marlins 20] defeated the tired [2
1 21 Cincinnati Reds 21] 5-3 . The [22 22 Marlins 22] won for
just the third time in nine games , taking advantage of a [23
23 Reds 23] ' team that has not had a day off since August
8th and was playing its fourth game in 43 hours . In [24 24 Lo
s Angeles 24] , [25 25 Tom Candiotti 25] allowed two runs in
seven innings and singled home the go-ahead run and [26 26 Mi
ke Piazza 26] and [27 27 Todd Hollandsworth 27] drove in two
runs apiece as the [28 28 Los Angeles Dodgers 28] defeated
the [29 29 New York Mets 29] 7-5 . [30 30 Candiotti 30]
(8-9) walked one , allowed five hits and struck out a
season-high eight batters for [31 31 Los Angeles 31] , which
has won 10 of its last 14 games . In [32 32 San Diego 32] , [3
3 33 Joey Hamilton 33] allowed two hits over seven innings
and [34 34 Rickey Henderson 34] hit his major league-record
69th leadoff homer as the [35 35 San Diego Padres 35]
defeated the [36 36 Philadelphia Phillies 36] 7-1 for their
fifth win in six games . [37 37 Hamilton 37] (12-7) won his
second straight start , allowing just a sixth-inning run and
a pair of singles . In [38 38 San Francisco 38] , [39 39 Pedro
Martinez 39] allowed two hits in eight innings and [40 40 Dav
id Segui 40] drove in two runs as the [41 41 Montreal Expos 41
] shut out the [42 42 San Francisco Giants 42] 3-0 for their
third straight win . [43 43 Martinez 43] (11-7) , who
lasted just 1-1/3 innings in his last start against [44 44 San
Diego 44] five days ago , pitched eight-plus innings ,
walking four and striking out 10 . In [45 45 Houston 45] , [46
46 Orlando Miller 46] 's two-run homer with one out in the

bottom of the ninth off [47 47 Todd Stottlemyre 47] gave the [48 48 Houston Astros 48] a 3-1 win over the [49 49 St. Louis Cardinals 49] and left the teams in a virtual tie for the lead in the NL Central division . [50 50 Shane Reynolds 50] (16-6) fired a five-hitter , walking one and striking out six .

1: 2140 Atlanta_Braves
2: 6654 Chicago_Cubs
3: 3138 Atlanta
4: 205905 Fred_McGriff
5: 2140 Atlanta_Braves
6: 6654 Chicago_Cubs
7: 205905 Fred_McGriff
8: 4258501 Brad_Clontz
9: 2140 Atlanta_Braves
10: 5399 Colorado
11: 9045614 Mark_Thompson_(baseball)
12: 932084 Ellis_Burks
13: 6669 Colorado_Rockies
14: 24270 Pittsburgh_Pirates
15: 933906 Vinny_Castilla
16: 419550 Dante_Bichette
17: 6669 Colorado_Rockies
18: 425563 Kevin_Brown_(right-handed_pitcher)
19: 6276814 Kurt_Abbott
20: 11278 Miami_Marlins
21: 6671 Cincinnati_Reds
22: 11278 Miami_Marlins
23: 6671 Cincinnati_Reds
24: 18110 Los_Angeles
25: 1581778 Tom_Candiotti
26: 182778 Mike_Piazza
27: 1813259 Todd_Hollandsworth
28: 18213 Los_Angeles_Dodgers
29: 21728 New_York_Mets
30: 1581778 Tom_Candiotti
31: 18213 Los_Angeles_Dodgers
32: 28504 San_Diego
33: 10388259 Joey_Hamilton
34: 65453 Rickey_Henderson
35: 28417 San_Diego_Padres
36: 23741 Philadelphia_Phillies
37: 10388259 Joey_Hamilton
38: 49728 San_Francisco
39: 24630 Pedro_Martinez
40: 2806454 David_Segui
41: 20153 Montreal_Expos
42: 28416 San_Francisco_Giants
43: 24630 Pedro_Martinez
44: 28417 San_Diego_Padres
45: 13774 Houston
46: 6085940 Orlando_Miller

```
47: 2572256 Todd_Stottlemyre
48: 13894 Houston_Astros
49: 23408844 St._Louis_Cardinals
50: 4526365 Shane_Reynolds
1: 11278 Miami_Marlins
    [winner: 18933066 Florida, pem_pos=0,
final_scores=-0.40426815, final_local_scores=-0.39950097,
pem=-0.15405534, similarity=-6.9944158]
    RECALL MISS: 11278 Miami_Marlins]
1: 2140 Atlanta_Braves
    [pem_pos=0, final_scores=0.5406319,
final_local_scores=0.48160356, pem=-0.28448334,
similarity=20.00879]
2: 6654 Chicago_Cubs
    [pem_pos=0, final_scores=0.57635695,
final_local_scores=0.5213201, pem=-0.2883288,
similarity=21.214922]
3: 3138 Atlanta
    [pem_pos=0, final_scores=0.5323344,
final_local_scores=0.59645957, pem=-0.23152597,
similarity=23.14833]
4: 205905 Fred_McGriff
    [pem_pos=0, final_scores=0.6506655,
final_local_scores=0.65789646, pem=0.0, similarity=23.722656]
5: 2140 Atlanta_Braves
    [pem_pos=0, final_scores=0.5289814,
final_local_scores=0.46887714, pem=-0.06390974,
similarity=18.429493]
6: 6654 Chicago_Cubs
    [pem_pos=0, final_scores=0.45692438,
final_local_scores=0.39085817, pem=-0.085107334,
similarity=16.216528]
7: 205905 Fred_McGriff
    [pem_pos=0, final_scores=0.8120635,
final_local_scores=0.8375341, pem=0.0, similarity=29.083403]
8: 4258501 Brad_Clontz
    [pem_pos=0, final_scores=0.6343409,
final_local_scores=0.71461594, pem=0.0, similarity=25.415277]
9: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=0.82878107,
final_local_scores=0.7963628, pem=-4.136474,
similarity=50.349556]
10: 5399 Colorado
    [pem_pos=0, final_scores=0.4548639,
final_local_scores=0.5593451, pem=-0.26926568,
similarity=22.245996]
11: 9045614 Mark_Thompson_(baseball)
    [pem_pos=4, final_scores=0.4096554,
final_local_scores=0.4689418, pem=-1.8939172,
similarity=28.383287]
12: 932084 Ellis_Burks
    [pem_pos=0, final_scores=0.9072774,
```

```
final_local_scores=0.9643205, pem=0.0, similarity=32.86696]
13: 6669 Colorado_Rockies
    [pem_pos=0, final_scores=0.45558763,
final_local_scores=0.390579, pem=-0.10158596,
similarity=16.29781]
14: 24270 Pittsburgh_Pirates
    [pem_pos=0, final_scores=0.541381,
final_local_scores=0.48323637, pem=-0.15187988,
similarity=19.336395]
15: 933906 Vinny_Castilla
    [pem_pos=0, final_scores=0.66066253,
final_local_scores=0.68481535, pem=0.0, similarity=24.525969]
16: 419550 Dante_Bichette
    [pem_pos=0, final_scores=0.63445705,
final_local_scores=0.65464437, pem=0.0, similarity=23.625607]
17: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.0114233,
final_local_scores=0.9977451, pem=-4.482383,
similarity=58.24031]
18: 425563 Kevin_Brown_(right-handed_pitcher)
    [pem_pos=0, final_scores=0.61809593,
final_local_scores=0.63052875, pem=-0.26213035,
similarity=24.331453]
19: 6276814 Kurt_Abbott
    [pem_pos=0, final_scores=0.68961895,
final_local_scores=0.75957036, pem=0.0, similarity=26.756807]
20: 11278 Miami_Marlins
    [pem_pos=0, final_scores=0.45650148,
final_local_scores=0.39320737, pem=0.0, similarity=15.823805]
21: 6671 Cincinnati_Reds
    [pem_pos=0, final_scores=0.5066938,
final_local_scores=0.4480719, pem=-0.106723964,
similarity=18.041452]
22: 11278 Miami_Marlins
    [pem_pos=0, final_scores=0.44723064,
final_local_scores=0.38308036, pem=-0.22471714,
similarity=16.743643]
23: 6671 Cincinnati_Reds
    [pem_pos=1, final_scores=0.4869172,
final_local_scores=0.42646897, pem=-1.0258155,
similarity=22.394941]
24: 18110 Los_Angeles
    [pem_pos=0, final_scores=0.57053995,
final_local_scores=0.6423487, pem=-0.100652285,
similarity=23.806042]
25: 1581778 Tom_Candiotti
    [pem_pos=0, final_scores=0.587764,
final_local_scores=0.6358462, pem=0.0, similarity=23.064632]
26: 182778 Mike_Piazza
    [pem_pos=0, final_scores=0.7895829,
final_local_scores=0.7970383, pem=0.0, similarity=27.874926]
27: 1813259 Todd_Hollandsworth
```

```
[pem_pos=0, final_scores=0.75148183,
final_local_scores=0.8036776, pem=0.0, similarity=28.073055]
28: 18213 Los_Angeles_Dodgers
    [pem_pos=0, final_scores=0.62753934,
final_local_scores=0.5730088, pem=-0.105150424,
similarity=21.761263]
29: 21728 New_York_Mets
    [pem_pos=0, final_scores=0.5476569,
final_local_scores=0.4890657, pem=-0.065782696,
similarity=19.042145]
30: 1581778 Tom_Candiotti
    [pem_pos=0, final_scores=0.72000027,
final_local_scores=0.78029424, pem=0.0, similarity=27.37525]
31: 18213 Los_Angeles_Dodgers
    [pem_pos=15, final_scores=0.71097356,
final_local_scores=0.66414803, pem=-4.7524934,
similarity=49.754013]
32: 28504 San_Diego
    [pem_pos=0, final_scores=0.5091349,
final_local_scores=0.5646225, pem=-0.18777876,
similarity=21.960344]
33: 10388259 Joey_Hamilton
    [pem_pos=0, final_scores=0.6272448,
final_local_scores=0.642429, pem=0.0, similarity=23.261076]
34: 65453 Rickey_Henderson
    [pem_pos=0, final_scores=0.64101034,
final_local_scores=0.64496714, pem=0.0, similarity=23.336819]
35: 28417 San_Diego_Padres
    [pem_pos=0, final_scores=0.46253276,
final_local_scores=0.39518237, pem=-0.05076447,
similarity=16.158808]
36: 23741 Philadelphia_Phillies
    [pem_pos=0, final_scores=0.47249717,
final_local_scores=0.40871733, pem=-0.09993234,
similarity=16.830101]
37: 10388259 Joey_Hamilton
    [pem_pos=0, final_scores=1.1260016,
final_local_scores=1.1872451, pem=0.0, similarity=39.519474]
38: 49728 San_Francisco
    [pem_pos=0, final_scores=0.51704454,
final_local_scores=0.58601177, pem=-0.09025385,
similarity=22.068289]
39: 24630 Pedro_Martinez
    [pem_pos=0, final_scores=0.45886296,
final_local_scores=0.44704503, pem=0.0, similarity=17.430428]
40: 2806454 David_Segui
    [pem_pos=0, final_scores=0.47874928,
final_local_scores=0.5233291, pem=0.0, similarity=19.706896]
41: 20153 Montreal_Expos
    [pem_pos=0, final_scores=0.37719327,
final_local_scores=0.30977005, pem=-0.14979224,
similarity=14.148462]
```

```
42: 28416 San_Francisco_Giants
    [pem_pos=0, final_scores=0.5614498,
final_local_scores=0.5046488, pem=-0.07570464,
similarity=19.561134]
43: 24630 Pedro_Martinez
    [pem_pos=0, final_scores=0.9908994,
final_local_scores=1.0314963, pem=0.0, similarity=34.871616]
44: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.7819373,
final_local_scores=0.7440833, pem=-3.9307637,
similarity=47.670746]
45: 13774 Houston
    [pem_pos=0, final_scores=0.6692268,
final_local_scores=0.73712844, pem=-0.3232624,
similarity=27.845047]
46: 6085940 Orlando_Miller
    [pem_pos=0, final_scores=0.52884406,
final_local_scores=0.6075549, pem=0.0, similarity=22.220364]
47: 2572256 Todd_Stottlemyre
    [pem_pos=0, final_scores=0.54296994,
final_local_scores=0.57706356, pem=0.0, similarity=21.31044]
48: 13894 Houston_Astros
    [pem_pos=0, final_scores=0.38680625,
final_local_scores=0.3112291, pem=-0.049621258,
similarity=13.647259]
49: 23408844 St._Louis_Cardinals
    [pem_pos=0, final_scores=0.4318456,
final_local_scores=0.36698872, pem=-0.2366796,
similarity=16.32849]
50: 4526365 Shane_Reynolds
    [pem_pos=0, final_scores=0.80692905,
final_local_scores=0.89709383, pem=0.0, similarity=30.860783]
```

Sample: b'352_BASEBALL'

```
BASEBALL - [1 1 MAJOR LEAGUE 1] RESULTS SATURDAY . [1 1 NEW Y
ORK 1] 1996-08-25 Results of [2 2 Major League 2] Baseball
games played on Saturday ( home team in CAPS ) : [3 3 National
League 3] [4 4 ATLANTA 4] 6 2 Chicago 5 [5 5 HOUSTON 5] 3 St
Louis 1 3 LOS ANGELES 7 [6 6 New York 6] 5 [7 7 Montreal 7] 3
[8 8 SAN FRANCISCO 8] 0 4 FLORIDA 5 [9 9 Cincinnati 9] 3 [10 1
0 COLORADO 10] 9 [11 11 Pittsburgh 11] 3 [12 12 SAN DIEGO 12]
7 [13 13 Philadelphia 13] 1 [14 14 American League 14] [15 15
BOSTON 15] 9 [16 16 Seattle 16] 5 [17 17 Milwaukee 17] 4 [18 1
8 CLEVELAND 18] 3 ( 10 innings ) [19 19 BALTIMORE 19] 5 5 Cal
ifornia 4 [20 20 Toronto 20] 9 6 CHICAGO 2 [21 21 NEW YORK 21]
5 [22 22 Oakland 22] 4 [23 23 KANSAS CITY 23] 9 [24 24 Detroit
24] 2 [25 25 MINNESOTA 25] 6 [26 26 Texas 26] 5
1: 38776 Major_League_Baseball
2: 38776 Major_League_Baseball
3: 64290 National_League
4: 2140 Atlanta_Braves
5: 13894 Houston_Astros
```

```
6: 4848143 New_York_Yankees
7: 20153 Montreal_Expos
8: 28416 San_Francisco_Giants
9: 6671 Cincinnati_Reds
10: 6669 Colorado_Rockies
11: 24270 Pittsburgh_Pirates
12: 28417 San_Diego_Padres
13: 23741 Philadelphia_Phillies
14: 64777 American_League
15: 4339 Boston_Red_Sox
16: 27660 Seattle_Mariners
17: 20152 Milwaukee_Brewers
18: 6652 Cleveland_Indians
19: 4340 Baltimore_Orioles
20: 30847 Toronto_Blue_Jays
21: 4848143 New_York_Yankees
22: 22523 Oakland_Athletics
23: 16883 Kansas_City_Royals
24: 8579 Detroit_Tigers
25: 20050 Minnesota_Twins
26: 30857 Texas_Rangers_(baseball)
1: 8210131 New_York
    [winner: 645042 New_York_City, pem_pos=1,
final_scores=0.24043559, final_local_scores=0.3038904,
pem=-1.7906687, similarity=22.896343]
    gt: 8210131 New_York, pem_pos=0,
final_scores=-0.049160212, final_local_scores=-0.0065287948,
pem=-0.2889045, similarity=5.465988 ]
2: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.57888347, final_local_scores=-0.65257114,
pem=-4.591283, similarity=9.5837965|
     RECALL MISS: 5945 Chicago_White_Sox]
3: 1360083 Los_Angeles_Angels_of_Anheim
    [winner: 18213 Los_Angeles_Dodgers, pem_pos=15,
final_scores=-0.1989251, final_local_scores=-0.33248222,
pem=-4.7524934, similarity=20.012579|
     RECALL MISS: 1360083 Los_Angeles_Angels_of_Anheim]
4: 11278 Miami_Marlins
    [winner: 18933066 Florida, pem_pos=0,
final_scores=-0.73292226, final_local_scores=-0.7540923,
pem=-0.15405534, similarity=-17.576128|
     RECALL MISS: 11278 Miami_Marlins]
5: 1360083 Los_Angeles_Angels_of_Anheim
    [winner: 7590031 California_Golden_Bears_football,
pem_pos=1, final_scores=-0.7705491,
final_local_scores=-0.8436187, pem=-4.4771376,
similarity=3.2618191|
     RECALL MISS: 1360083 Los_Angeles_Angels_of_Anheim]
6: 5945 Chicago_White_Sox
    [winner: 994449 Chicago_(album), pem_pos=0,
final_scores=-0.7714726, final_local_scores=-0.7150332,
```

```
pem=0.0, similarity=-17.248302|
    RECALL MISS: 5945 Chicago_White_Sox]
1: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.5598782,
final_local_scores=0.5069253, pem=-0.47792417,
similarity=21.8164]
2: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.55802274,
final_local_scores=0.5048985, pem=-0.47792417,
similarity=21.755917]
3: 64290 National_League
    [pem_pos=0, final_scores=0.55874974,
final_local_scores=0.5232961, pem=-0.07431644,
similarity=20.110058]
4: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=1.1103439,
final_local_scores=1.0969417, pem=-4.136474,
similarity=59.31943]
5: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.1775692,
final_local_scores=1.172576, pem=-4.650927,
similarity=64.37418]
6: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.53244716,
final_local_scores=0.4686727, pem=-5.0692368,
similarity=45.64314]
7: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.8408175,
final_local_scores=0.8096782, pem=-4.5098605,
similarity=52.777447]
8: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.8755985,
final_local_scores=0.8407535, pem=-4.409061,
similarity=53.15663]
9: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.1558298,
final_local_scores=1.1470613, pem=-3.771247,
similarity=58.82894]
10: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.1013892,
final_local_scores=1.0973196, pem=-4.482383,
similarity=61.21181]
11: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.1711361,
final_local_scores=1.1619709, pem=-4.03702,
similarity=60.71918]
12: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.97498494,
final_local_scores=0.9553743, pem=-3.9307637,
similarity=53.97609]
13: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=0.9992179,
```

```
final_local_scores=0.97622186, pem=-4.5189333,
similarity=57.79678]
14: 64777 American_League
    [pem_pos=0, final_scores=0.55221695,
final_local_scores=0.52135026, pem=0.0, similarity=19.647844]
15: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.1101665,
final_local_scores=1.1007609, pem=-4.418931,
similarity=60.969448]
16: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.5409749,
final_local_scores=0.48571712, pem=-3.8567436,
similarity=39.558052]
17: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.5190443,
final_local_scores=0.45549852, pem=-3.40205,
similarity=36.183575]
18: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.0792394,
final_local_scores=1.0614455, pem=-4.191953,
similarity=58.56186]
19: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.3041918,
final_local_scores=1.3077791, pem=-4.046561,
similarity=65.12228]
20: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.9212641,
final_local_scores=0.9005032, pem=-4.422637,
similarity=55.01351]
21: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.551589,
final_local_scores=0.48958224, pem=-5.0692368,
similarity=46.26712]
22: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.0635357,
final_local_scores=1.0460484, pem=-3.9188845,
similarity=56.617386]
23: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.8771797,
final_local_scores=0.84683406, pem=-3.5915844,
similarity=48.892525]
24: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.0173894,
final_local_scores=0.9947558, pem=-4.1401463,
similarity=56.28997]
25: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=1.1262636,
final_local_scores=1.1156948, pem=-4.2906566,
similarity=60.71753]
26: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=0.9514993,
final_local_scores=0.9274656, pem=-4.928425,
```

```
similarity=58.568676]
1: 645042 New_York_City
    [pem_pos=1, final_scores=0.24043559,
final_local_scores=0.3038904, pem=-1.7906687,
similarity=22.896343]

Sample: b'353_SOCCER'
SOCCER - CHAMPIONS [1 1 PORTO 1] KICK OFF SEASON WITH A
DRAW . [2 2 LISBON 2] 1996-08-25 [3 3 Portuguese 3] champions
[4 4 Porto 4] kicked off the season with a disappointing 2-2
home draw against Setubal and were lucky to squeeze in an
equaliser in extra time . [5 5 Porto 5] , who are fighting to
take their third consecutive title this season , were 2-0
down until the 86th minute when a header by [6 6 Mario Jardel
6] found the net after a string of missed opportunities ,
including a penalty taken by top league scorer Domingos
Oliveira in the 60th minute . Domingos redeemed himself by
netting the equaliser just into extra time . Setubal , who
put on a skilful counter-attack throughout the game , opened
the scoring 16 minutes into the match when an unmarked [7 7 Ch
iquinho Conde 7] shot around [8 8 Porto 8] 's new [9 9 Polish
9] keeper Andrezej Wozniak . [10 10 Conde 10] scored his
second in the 70th minute . [11 11 Benfica 11] , also playing
their first game of the season at home , were held to a 1-1
draw by northern side [12 12 Braga 12] despite the fact that
the visitors were reduced to 10 men in the 54th minute after
Rodrigo Carneiro was sent off for a second bookable offence .
[13 13 Benfica 13] dominated the game but their lack of a
first-class striker was apparent throughout and in the 30th
minute they lost key [14 14 Brazilian 14] midfielder [15 15 Va
ldo 15] who suffered a light knee injury . He was substituted
by [16 16 Paulao 16] . [17 17 Benfica 17] finally opened the
scoring in the 81st minute with a penalty taken by Helder
after Luis Baltasar tripped up captain [18 18 Joao Pinto 18]
under the referee 's nose . [19 19 Braga 19] defender
Idalecio gave his team their equaliser seven minutes from the
final whistle with a header into the back of the net .
1: 172589 F.C._Porto
2: 18091 Lisbon
3: 23033 Portugal
4: 172589 F.C._Porto
5: 172589 F.C._Porto
6: 766470 Mário_Jardel
7: 8295532 Chiquinho_Conde
8: 172589 F.C._Porto
9: 22936 Poland
10: 8295532 Chiquinho_Conde
11: 574076 S.L._Benfica
12: 990671 S.C._Braga
13: 574076 S.L._Benfica
14: 3383 Brazil
15: 3860806 Valdo_Filho
```

```
16: 18884152 Paulo_Roberto_do_Carmo
17: 574076 S.L._Benfica
18: 1445352 João_Domingos_Pinto
19: 990671 S.C._Braga
1: 172589 F.C._Porto
    [pem_pos=1, final_scores=0.87251,
final_local_scores=0.8537172, pem=-1.1013846,
similarity=35.55584]
2: 18091 Lisbon
    [pem_pos=0, final_scores=0.52002853,
final_local_scores=0.5468565, pem=-0.09749677,
similarity=20.939205]
3: 23033 Portugal
    [pem_pos=1, final_scores=0.46565622,
final_local_scores=0.48593426, pem=-1.127839,
similarity=24.724325]
4: 172589 F.C._Porto
    [pem_pos=1, final_scores=0.6659717,
final_local_scores=0.6281055, pem=-1.1013846,
similarity=28.823135]
5: 172589 F.C._Porto
    [pem_pos=1, final_scores=0.6552731,
final_local_scores=0.61641896, pem=-1.1013846,
similarity=28.474386]
6: 766470 Mário_Jardel
    [pem_pos=0, final_scores=0.53949577,
final_local_scores=0.56700534, pem=0.0, similarity=21.010283]
7: 8295532 Chiquinho_Conde
    [pem_pos=0, final_scores=0.497037,
final_local_scores=0.5279189, pem=0.0, similarity=19.843864]
8: 172589 F.C._Porto
    [pem_pos=1, final_scores=0.5896396,
final_local_scores=0.5447243, pem=-1.1013846,
similarity=26.334873]
9: 22936 Poland
    [pem_pos=0, final_scores=0.46169007,
final_local_scores=0.54941255, pem=-0.81518245,
similarity=24.918367]
10: 8295532 Chiquinho_Conde
    [pem_pos=0, final_scores=0.3669886,
final_local_scores=0.3858608, pem=0.0, similarity=15.6045685]
11: 574076 S.L._Benfica
    [pem_pos=0, final_scores=0.5069675,
final_local_scores=0.44896752, pem=0.0, similarity=17.487799]
12: 990671 S.C._Braga
    [pem_pos=1, final_scores=0.49453694,
final_local_scores=0.44517428, pem=-1.3993924,
similarity=24.984715]
13: 574076 S.L._Benfica
    [pem_pos=0, final_scores=0.5102608,
final_local_scores=0.45256495, pem=0.0, similarity=17.595154]
14: 3383 Brazil
```

```
[pem_pos=0, final_scores=0.65798604,
final_local_scores=0.7284168, pem=-0.2733083,
similarity=27.313416]
15: 3860806 Valdo_Filho
    [pem_pos=0, final_scores=0.42831093,
final_local_scores=0.47648233, pem=-0.45246792,
similarity=20.769487]
16: 18884152 Paulo_Roberto_do_Carmo
    [pem_pos=0, final_scores=0.40475267,
final_local_scores=0.4876873, pem=0.0, similarity=18.643274]
17: 574076 S.L._Benfica
    [pem_pos=0, final_scores=0.44271755,
final_local_scores=0.37878418, pem=0.0, similarity=15.393389]
18: 1445352 João_Domingos_Pinto
    [pem_pos=0, final_scores=0.32622498,
final_local_scores=0.35018402, pem=0.0, similarity=14.539903]
19: 990671 S.C._Braga
    [pem_pos=1, final_scores=0.5120514,
final_local_scores=0.46430618, pem=-1.3993924,
similarity=25.555647]

Sample: b'354_SOCCER'
SOCCER - [1 1 PORTUGUESE 1] FIRST DIVISION RESULTS . [2 2 LI
SBON 2] 1996-08-25 Results of [3 3 Portuguese 3] first
division soccer matches on Sunday : [4 4 FC Porto 4] 2
Setubal 2 [5 5 Benfica 5] 1 [6 6 Braga 6] 1 [7 7 Guimaraes 7]
4 [8 8 Gil Vicente 8] 2
1: 23033 Portugal
2: 18091 Lisbon
3: 576328 Portugal_national_football_team
4: 172589 F.C._Porto
5: 574076 S.L._Benfica
6: 990671 S.C._Braga
7: 1127372 Vitória_S.C.
8: 1629212 Gil_Vicente_F.C.
1: 23033 Portugal
    [pem_pos=1, final_scores=0.27431613,
final_local_scores=0.28445035, pem=-1.127839,
similarity=18.711643]
2: 18091 Lisbon
    [pem_pos=0, final_scores=0.6118812,
final_local_scores=0.64938116, pem=-0.09749677,
similarity=23.998745]
3: 576328 Portugal_national_football_team
    [pem_pos=14, final_scores=0.14663516,
final_local_scores=0.07996851, pem=-3.9943988,
similarity=28.198294]
4: 172589 F.C._Porto
    [pem_pos=0, final_scores=0.45908213,
final_local_scores=0.40992296, pem=0.0, similarity=16.322632]
5: 574076 S.L._Benfica
    [pem_pos=0, final_scores=0.44115227,
```

```
final_local_scores=0.38406378, pem=0.0, similarity=15.550941]
6: 990671 S.C._Braga
    [pem_pos=1, final_scores=0.43166208,
final_local_scores=0.37917048, pem=-1.3993924,
similarity=23.01503]
7: 1127372 Vitória_S.C.
    [pem_pos=1, final_scores=0.4007966,
final_local_scores=0.34527576, pem=-1.0001721,
similarity=19.832521]
8: 1629212 Gil_Vicente_F.C.
    [pem_pos=0, final_scores=0.46728605,
final_local_scores=0.44104242, pem=0.0, similarity=17.251299]
```

Sample: b'355_SOCCER'

SOCCER - [1 1 FIORENTINA 1] WIN WITH [2 2 BATISTUTA 2]
DOUBLE . [3 3 MILAN 3] 1996-08-25 [4 4 Argentine 4] striker [5
5 Gabriel Batistuta 5] gave [6 6 Fiorentina 6] the perfect
70th birthday present on Sunday with two goals that gave the [
7 7 Italian Cup 7] winners a 2-1 Supercup victory over [8 8 se
rie A 8] champions [9 9 Milan 9] . The victory , coming on
the eve of the founding of the [10 10 Florence 10] club in
1926 , also marked the first time since the pre-season trophy
between the 1 Cup winners and league champions was started in
1988 that the 2 Cup winners had won . [11 11 Batistuta 11]
gave [12 12 Fiorentina 12] the lead in the 11th minute . [13 1
3 Sweden 13] 's [14 14 Stefan Schwarz 14] picked him out with
a lob to the edge of the box and [15 15 Batistuta 15] did the
rest , chipping veteran defender [16 16 Franco Baresi 16] and
scoring at the near post . [17 17 Montenegrin 17] midfielder [
18 18 Dejan Savicevic 18] equalised for the home team ,
weaving past a defender , checking and firing in a left-
footed shot in the 21st minute that gave young [19 19 Fiorent
ina 19] goalkeeper [20 20 Francesco Toldo 20] little chance .
The score stayed level to the final minutes but with a
penalty shoot out looming , [21 21 Batistuta 21] took
charge . [22 22 French 22] international midfielder [23 23 Ma
rcel Desailly 23] fouled the [24 24 Argentine 24] , whose
coach at [25 25 Boca Juniors 25] before he joined [26 26 Fior
entina 26] in 1991 was new [27 27 Milan 27] coach [28 28 Oscar
Tabarez 28] , and [29 29 Batistuta 29] rammed home the free
kick from 30 metres out . The 83rd minute shot , curling over
the defence and dipping in under the bar from the striker
dubbed [30 30 Batigol 30] by adoring [31 31 Fiorentina 31]
fans , was just reward for [32 32 Fiorentina 32] who looked a
far more impressive team . [33 33 Milan 33] 's player of the
year [34 34 George Weah 34] missed a good first half
opportunity but otherwise looked a little rusty while [35 35 I
talian 35] team mate [36 36 Roberto Baggio 36] did not play
due to injury . New [37 37 Dutch 37] signing [38 38 Edgar Dav
ids 38] came on late in the second half as a [39 39 Milan 39]
substitute but made little impact . The league season starts
on September 8 .

```
1: 3165 ACF_Fiorentina
2: 557154 Gabriel_Batistuta
3: 36511 Milan
4: 18951905 Argentina
5: 557154 Gabriel_Batistuta
6: 3165 ACF_Fiorentina
7: 1281200 Coppa_Italia
8: 15496 Serie_A
9: 18940588 A.C._Milan
10: 11525 Florence
11: 557154 Gabriel_Batistuta
12: 3165 ACF_Fiorentina
13: 5058739 Sweden
14: 1896409 Stefan_Schwarz
15: 557154 Gabriel_Batistuta
16: 10855 Franco_Baresi
17: 592964 Montenegrins
18: 3307429 Dejan_Savićević
19: 3165 ACF_Fiorentina
20: 1372000 Francesco_Toldo
21: 557154 Gabriel_Batistuta
22: 5843419 France
23: 413244 Marcel_Desailly
24: 18951905 Argentina
25: 401877 Boca_Juniors
26: 3165 ACF_Fiorentina
27: 18940588 A.C._Milan
28: 9158864 Óscar_Tabárez
29: 557154 Gabriel_Batistuta
30: 557154 Gabriel_Batistuta
31: 3165 ACF_Fiorentina
32: 3165 ACF_Fiorentina
33: 18940588 A.C._Milan
34: 616629 George_Weah
35: 362466 Italy_national_football_team
36: 25624 Roberto_Baggio
37: 21148 Netherlands
38: 252130 Edgar_Davids
39: 18940588 A.C._Milan
1: 1281200 Coppa_Italia
    [winner: 11237 FA_Cup, pem_pos=6,
final_scores=-0.22116746, final_local_scores=-0.2479513,
pem=-3.1209865, similarity=13.662761|
     RECALL MISS: 1281200 Coppa_Italia]
2: 1281200 Coppa_Italia
    [winner: 11237 FA_Cup, pem_pos=6,
final_scores=-0.28716207, final_local_scores=-0.32004037,
pem=-3.1209865, similarity=11.511479|
     RECALL MISS: 1281200 Coppa_Italia]
1: 3165 ACF_Fiorentina
    [pem_pos=0, final_scores=0.73229915,
final_local_scores=0.70489854, pem=0.0, similarity=25.125292]
```

```
2: 557154 Gabriel_Batistuta
    [pem_pos=0, final_scores=0.74440664,
final_local_scores=0.7408985, pem=0.0, similarity=26.199602]
3: 36511 Milan
    [pem_pos=0, final_scores=0.39240956,
final_local_scores=0.42909825, pem=-0.3015942,
similarity=18.534977]
4: 18951905 Argentina
    [pem_pos=0, final_scores=0.5572728,
final_local_scores=0.6226891, pem=-0.18885723,
similarity=23.699034]
5: 557154 Gabriel_Batistuta
    [pem_pos=0, final_scores=0.88575,
final_local_scores=0.8952946, pem=0.0, similarity=30.80709]
6: 3165 ACF_Fiorentina
    [pem_pos=0, final_scores=0.5228695,
final_local_scores=0.4761284, pem=0.0, similarity=18.298334]
7: 1281200 Coppa_Italia
    [pem_pos=0, final_scores=0.275562,
final_local_scores=0.23616737, pem=0.0, similarity=11.137419]
8: 15496 Serie_A
    [pem_pos=0, final_scores=0.4953174,
final_local_scores=0.45080262, pem=-0.1387667,
similarity=18.297197]
9: 18940588 A.C._Milan
    [pem_pos=1, final_scores=0.74530786,
final_local_scores=0.7117576, pem=-1.9620993,
similarity=36.00018]
10: 11525 Florence
    [pem_pos=0, final_scores=0.5192578,
final_local_scores=0.5799142, pem=-0.32229787,
similarity=23.148216]
11: 557154 Gabriel_Batistuta
    [pem_pos=0, final_scores=0.5535899,
final_local_scores=0.5324601, pem=0.0, similarity=19.979383]
12: 3165 ACF_Fiorentina
    [pem_pos=0, final_scores=0.47216552,
final_local_scores=0.42074203, pem=0.0, similarity=16.645494]
13: 5058739 Sweden
    [pem_pos=0, final_scores=0.6519709,
final_local_scores=0.7436791, pem=-0.2135508,
similarity=27.443903]
14: 1896409 Stefan_Schwarz
    [pem_pos=0, final_scores=0.66325885,
final_local_scores=0.72537124, pem=0.0, similarity=25.736238]
15: 557154 Gabriel_Batistuta
    [pem_pos=0, final_scores=0.77073294,
final_local_scores=0.76965594, pem=0.0, similarity=27.057781]
16: 10855 Franco_Baresi
    [pem_pos=0, final_scores=0.65538204,
final_local_scores=0.7029799, pem=0.0, similarity=25.068035]
17: 592964 Montenegrins
```

```
[pem_pos=2, final_scores=0.11405398,
final_local_scores=0.15335783, pem=-1.1853291,
similarity=15.11222]
18: 3307429 Dejan_Savićević
    [pem_pos=0, final_scores=1.0173974,
final_local_scores=1.0830159, pem=0.0, similarity=36.409065]
19: 3165 ACF_Fiorentina
    [pem_pos=0, final_scores=0.54684347,
final_local_scores=0.5023163, pem=0.0, similarity=19.079834]
20: 1372000 Francesco_Toldo
    [pem_pos=0, final_scores=0.6971442,
final_local_scores=0.719472, pem=0.0, similarity=25.560192]
21: 557154 Gabriel_Batistuta
    [pem_pos=0, final_scores=0.7316046,
final_local_scores=0.72691417, pem=0.0, similarity=25.782282]
22: 5843419 France
    [pem_pos=0, final_scores=0.6151224,
final_local_scores=0.6963545, pem=-0.80223274,
similarity=29.232986]
23: 413244 Marcel_Desailly
    [pem_pos=0, final_scores=0.7312445,
final_local_scores=0.7771495, pem=0.0, similarity=27.281404]
24: 18951905 Argentina
    [pem_pos=0, final_scores=0.46381837,
final_local_scores=0.5206044, pem=-0.18885723,
similarity=20.652622]
25: 401877 Boca_Juniors
    [pem_pos=0, final_scores=0.38086194,
final_local_scores=0.35288945, pem=0.0, similarity=14.620638]
26: 3165 ACF_Fiorentina
    [pem_pos=0, final_scores=0.4480496,
final_local_scores=0.394399, pem=0.0, similarity=15.859365]
27: 18940588 A.C._Milan
    [pem_pos=1, final_scores=0.6999897,
final_local_scores=0.6622544, pem=-1.9620993,
similarity=34.522907]
28: 9158864 Óscar_Tabárez
    [pem_pos=0, final_scores=0.6090656,
final_local_scores=0.6605266, pem=0.0, similarity=23.801144]
29: 557154 Gabriel_Batistuta
    [pem_pos=0, final_scores=0.6838671,
final_local_scores=0.6747682, pem=0.0, similarity=24.226143]
30: 557154 Gabriel_Batistuta
    [pem_pos=0, final_scores=0.6590324,
final_local_scores=0.64764005, pem=0.0, similarity=23.416584]
31: 3165 ACF_Fiorentina
    [pem_pos=0, final_scores=0.59332365,
final_local_scores=0.5530889, pem=0.0, similarity=20.59499]
32: 3165 ACF_Fiorentina
    [pem_pos=0, final_scores=0.54633397,
final_local_scores=0.50175977, pem=0.0, similarity=19.063225]
33: 18940588 A.C._Milan
```

```
[pem_pos=1, final_scores=0.7724188,
final_local_scores=0.7413723, pem=-1.9620993,
similarity=36.88394]
34: 616629 George_Weah
    [pem_pos=0, final_scores=0.91226035,
final_local_scores=0.9690638, pem=0.0, similarity=33.00851]
35: 362466 Italy_national_football_team
    [pem_pos=10, final_scores=0.42738807,
final_local_scores=0.38727015, pem=-4.774446,
similarity=41.610806]
36: 25624 Roberto_Baggio
    [pem_pos=0, final_scores=0.6103069,
final_local_scores=0.60182333, pem=0.0, similarity=22.049322]
37: 21148 Netherlands
    [pem_pos=0, final_scores=0.40187156,
final_local_scores=0.45522165, pem=-0.7943452,
similarity=21.994207]
38: 252130 Edgar_Davids
    [pem_pos=0, final_scores=0.70357376,
final_local_scores=0.73396635, pem=0.0, similarity=25.992733]
39: 18940588 A.C._Milan
    [pem_pos=1, final_scores=0.69351035,
final_local_scores=0.6551767, pem=-1.9620993,
similarity=34.311695]
```

Sample: b'356_SOCCER'

```
SOCCER - [1 1 Fiorentina 1] BEAT [2 2 Milan 2] IN [3 3 Italian SuperCup 3] . [4 4 Milan 4] 1996-08-25 [1 Italian Cup 1] winners [5 5 Fiorentina 5] beat league champions [6 6 Milan 6] 2-1 ( halftime 1-1 ) in the pre-season Supercoppa ( SuperCup ) in [7 7 Milan 7] on Sunday : Scorers : [8 8 Fiorentina 8] - [9 9 Gabriel Batistuta 9] ( 11th , 83rd ) [10 10 Milan 10] - [11 11 Dejan Savicevic 11] ( 21st ) Attendance : 29,582
```

1: 3165 ACF_Fiorentina
2: 18940588 A.C._Milan
3: 3035684 Supercoppa_Italiana
4: 36511 Milan
5: 3165 ACF_Fiorentina
6: 18940588 A.C._Milan
7: 36511 Milan
8: 3165 ACF_Fiorentina
9: 557154 Gabriel_Batistuta
10: 18940588 A.C._Milan
11: 3307429 Dejan_Savićević
1: 1281200 Coppa_Italia
 [winner: 1281200 Coppa_Italia, pem_pos=0,
final_scores=0.10826786, final_local_scores=0.040289298,
pem=0.0, similarity=5.292026]
 gt: 1281200 Coppa_Italia, pem_pos=0,
final_scores=0.10826786, final_local_scores=0.040289298,
pem=0.0, similarity=5.292026]

```

1: 3165 ACF_Fiorentina
    [pem_pos=0, final_scores=0.81033546,
final_local_scores=0.7805432, pem=0.0, similarity=27.382679]
2: 18940588 A.C._Milan
    [pem_pos=1, final_scores=0.58279914,
final_local_scores=0.5265798, pem=-1.9620993,
similarity=30.474108]
3: 3035684 Supercoppa_Italiana
    [pem_pos=0, final_scores=0.4172781,
final_local_scores=0.4098572, pem=0.0, similarity=16.32067]
4: 36511 Milan
    [pem_pos=0, final_scores=0.45290637,
final_local_scores=0.48775643, pem=-0.3015942,
similarity=20.285454]
5: 3165 ACF_Fiorentina
    [pem_pos=0, final_scores=0.5193234,
final_local_scores=0.4626568, pem=0.0, similarity=17.896315]
6: 18940588 A.C._Milan
    [pem_pos=1, final_scores=0.7060496,
final_local_scores=0.66121227, pem=-1.9620993,
similarity=34.491806]
7: 36511 Milan
    [pem_pos=0, final_scores=0.49437183,
final_local_scores=0.53305113, pem=-0.3015942,
similarity=21.637138]
8: 3165 ACF_Fiorentina
    [pem_pos=0, final_scores=0.60297316,
final_local_scores=0.55403143, pem=0.0, similarity=20.623116]
9: 557154 Gabriel_Batistuta
    [pem_pos=0, final_scores=0.86273426,
final_local_scores=0.9138927, pem=0.0, similarity=31.362095]
10: 18940588 A.C._Milan
    [pem_pos=1, final_scores=0.62294096,
final_local_scores=0.57042867, pem=-1.9620993,
similarity=31.782644]
11: 3307429 Dejan_Savićević
    [pem_pos=0, final_scores=0.78437346,
final_local_scores=0.8407699, pem=0.0, similarity=29.179964]
1: 1281200 Coppa_Italia
    [pem_pos=0, final_scores=0.10826786,
final_local_scores=0.040289298, pem=0.0, similarity=5.292026]

```

Sample: b'357_SOCCER'

```

SOCCER - [1 1 NORWAY 1] ELITE DIVISION RESULTS / STANDINGS .
[2 2 OSL0 2] 1996-08-25 Results of [3 3 Norwegian 3] elite
division soccer matches at the weekend : Tromso 2 [4 4 Kongsv
inger 4] 1 [5 5 Valerenga 5] 3 [6 6 Skeid 6] 0 [7 7 Stabaek 7]
4 Stromsgodset 0 [8 8 Molde 8] 1 Bodo / Glimt 2 [9 9 Viking 9]
1 [10 10 Moss 10] 0 [11 11 Brann 11] 7 [12 12 Start 12] 1 [13
13 Rosenborg 13] 7 [1 Lillestrom 1] 2 Standings after weekend
matches ( tabulate under played , won , drawn , lost , goals
for , against , points ) : [14 14 Rosenborg 14] 20 14 4 2 68

```

```
21 46 [2 Lillestrom 2] 19 9 5 5 38 29 32 [15 15 Skeid 15] 19  
10 2 7 29 30 32 [16 16 Stabaek 16] 20 7 8 5 41 34 29 [17 17 Br  
ann 17] 19 8 5 6 40 37 29 Tromso 20 8 5 7 34 33 29 [18 18 Vik  
ing 18] 20 7 7 6 33 24 28 [19 19 Molde 19] 19 8 3 8 36 25 27  
Bodo / Glimt 20 7 4 9 33 41 25 [20 20 Kongsvinger 20] 20 7 4  
9 26 38 25 Stromsgodset 20 7 4 9 27 40 25 [21 21 Vallerenga 21]  
20 6 6 8 26 32 24 [22 22 Moss 22] 20 4 6 10 23 40 18 [23 23 St  
art 23] 20 3 3 14 26 56 12  
1: 21241 Norway  
2: 22309 Oslo  
3: 21241 Norway  
4: 1171960 Kongsvinger_IL_Toppfotball  
5: 337678 Vålerenga_Fotball  
6: 1029557 Skeid_Fotball  
7: 318926 Stabæk_Fotball  
8: 176800 Molde_FK  
9: 214192 Viking_FK  
10: 181412 Moss_FK  
11: 192068 SK_Brann  
12: 1214850 IK_Start  
13: 86945 Rosenborg_BK  
14: 86945 Rosenborg_BK  
15: 1029557 Skeid_Fotball  
16: 318926 Stabæk_Fotball  
17: 192068 SK_Brann  
18: 214192 Viking_FK  
19: 176800 Molde_FK  
20: 1171960 Kongsvinger_IL_Toppfotball  
21: 337678 Vålerenga_Fotball  
22: 181412 Moss_FK  
23: 1214850 IK_Start  
1: 21241 Norway  
    [pem_pos=0, final_scores=0.6926706,  
final_local_scores=0.7655478, pem=-0.22668563,  
similarity=28.167938]  
2: 22309 Oslo  
    [pem_pos=0, final_scores=0.5264103,  
final_local_scores=0.5845112, pem=-0.035297178,  
similarity=21.724644]  
3: 21241 Norway  
    [pem_pos=0, final_scores=0.650393,  
final_local_scores=0.719366, pem=-0.605967,  
similarity=28.852373]  
4: 1171960 Kongsvinger_IL_Toppfotball  
    [pem_pos=1, final_scores=0.40866464,  
final_local_scores=0.36770773, pem=-1.1320317,  
similarity=21.21901]  
5: 337678 Vålerenga_Fotball  
    [pem_pos=0, final_scores=0.5053307,  
final_local_scores=0.45493644, pem=0.0, similarity=17.665924]  
6: 1029557 Skeid_Fotball  
    [pem_pos=0, final_scores=0.32707256,
```

```
final_local_scores=0.2852251, pem=0.0, similarity=12.601398]
7: 318926 Stabæk_Fotball
    [pem_pos=0, final_scores=0.41930377,
final_local_scores=0.373226, pem=0.0, similarity=15.227521]
8: 176800 Molde_FK
    [pem_pos=1, final_scores=0.70394164,
final_local_scores=0.6716818, pem=-0.5498513,
similarity=27.124214]
9: 214192 Viking_FK
    [pem_pos=1, final_scores=0.6460832,
final_local_scores=0.6053898, pem=-2.474389,
similarity=35.611866]
10: 181412 Moss_FK
    [pem_pos=2, final_scores=0.7206674,
final_local_scores=0.7154978, pem=-2.313225,
similarity=38.02127]
11: 192068 SK_Brann
    [pem_pos=0, final_scores=0.83041435,
final_local_scores=0.8053066, pem=-0.031611566,
similarity=28.293577]
12: 1214850 IK_Start
    [pem_pos=0, final_scores=1.0855858,
final_local_scores=1.0951409, pem=-0.5350669,
similarity=39.680683]
13: 86945 Rosenborg_BK
    [pem_pos=0, final_scores=0.4161088,
final_local_scores=0.35593715, pem=0.0, similarity=14.711588]
14: 86945 Rosenborg_BK
    [pem_pos=0, final_scores=0.39471167,
final_local_scores=0.33256406, pem=0.0, similarity=14.014088]
15: 1029557 Skeid_Fotball
    [pem_pos=0, final_scores=0.39294314,
final_local_scores=0.35717872, pem=0.0, similarity=14.748638]
16: 318926 Stabæk_Fotball
    [pem_pos=0, final_scores=0.42468673,
final_local_scores=0.37910604, pem=0.0, similarity=15.402993]
17: 192068 SK_Brann
    [pem_pos=0, final_scores=0.69112074,
final_local_scores=0.65314955, pem=-0.031611566,
similarity=23.752907]
18: 214192 Viking_FK
    [pem_pos=1, final_scores=0.68823814,
final_local_scores=0.6514377, pem=-2.474389,
similarity=36.986027]
19: 176800 Molde_FK
    [pem_pos=1, final_scores=0.61715394,
final_local_scores=0.57687944, pem=-0.5498513,
similarity=24.295122]
20: 1171960 Kongsvinger_IL_Toppfotball
    [pem_pos=1, final_scores=0.44466478,
final_local_scores=0.40703243, pem=-1.1320317,
similarity=22.392536]
```

```
21: 337678 Vålerenga_Fotball
    [pem_pos=0, final_scores=0.4710086,
final_local_scores=0.41744477, pem=0.0, similarity=16.547098]
22: 181412 Moss_FK
    [pem_pos=2, final_scores=0.63134015,
final_local_scores=0.6179213, pem=-2.313225,
similarity=35.109394]
23: 1214850 IK_Start
    [pem_pos=0, final_scores=0.9064551,
final_local_scores=0.89946777, pem=-0.5350669,
similarity=33.841404]
1: 87161 Lillestrøm_SK
    [pem_pos=1, final_scores=0.03901641,
final_local_scores=-0.049182013, pem=-1.0116009,
similarity=8.123266]
2: 87161 Lillestrøm_SK
    [pem_pos=1, final_scores=0.044753477,
final_local_scores=-0.04291515, pem=-1.0116009,
similarity=8.310282]
```

Sample: b'358_SOCCER'

```
SOCCER - SUMMARY OF [1 1 GERMAN 1] FIRST DIVISION MATCH . [2
2 BONN 2] 1996-08-25 Summary of [3 3 German 3] first division
match played on Sunday : MSV Duisberg 0 [4 4 Bayern Munich 4]
4 ( [5 5 Klinsmann 5] 15th , Zieger 24th and 90th , Witechek
59th ) . Halftime 0-2 . Attendance : 30,000 .
1: 11867 Germany
2: 3295 Bonn
3: 11867 Germany
4: 172326 FC_Bayern_Munich
5: 525141 Jürgen_Klinsmann
1: 11867 Germany
    [pem_pos=1, final_scores=0.84952,
final_local_scores=0.8509781, pem=-1.0465823,
similarity=35.176075]
2: 3295 Bonn
    [pem_pos=0, final_scores=0.46650922,
final_local_scores=0.4699812, pem=0.0, similarity=18.11489]
3: 11867 Germany
    [pem_pos=1, final_scores=0.7225395,
final_local_scores=0.71227115, pem=-1.0465823,
similarity=31.036783]
4: 172326 FC_Bayern_Munich
    [pem_pos=0, final_scores=0.3353706,
final_local_scores=0.35036805, pem=0.0, similarity=14.545395]
5: 525141 Jürgen_Klinsmann
    [pem_pos=0, final_scores=0.3924294,
final_local_scores=0.44075572, pem=0.0, similarity=17.242743]
```

Sample: b'359_SOCCER'

```
SOCCER - RESULT OF [1 1 GERMAN 1] FIRST DIVISION MATCH . [2 2
BONN 2] 1996-08-25 Result of [3 3 German 3] first division
```

soccer match on Sunday : MSV Duisberg 0 [4 4 Bayern Munich 4]
4 [5 5 Bundesliga 5] standings after Sunday 's game
(tabulate under played , won , drawn , lost , goals for ,
goals against , points) : [6 6 Cologne 6] 3 3 0 0 7 1 9 [7 7
Bayern Munich 7] 3 2 1 0 7 2 7 [8 8 VfB Stuttgart 8] 2 2 0 0
6 1 6 [9 9 Borussia Dortmund 9] 3 2 0 1 9 5 6 [10 10 Hamburg 1
0] 3 2 0 1 7 3 6 [11 11 Bayer Leverkusen 11] 3 2 0 1 7 4 6 [12
12 VfL Bochum 12] 3 1 2 0 3 2 5 [13 13 Karlsruhe 13] 2 1 1 0
5 3 4 [14 14 St Pauli 14] 3 1 1 1 7 7 4 [15 15 1860 Munich 15]
3 1 0 2 3 5 3 [16 16 Freiburg 16] 3 1 0 2 5 10 3 [17 17 Fortu
na Duesseldorf 17] 3 1 0 2 1 7 3 [18 18 Hansa Rostock 18] 3 0
2 1 3 4 2 [19 19 Arminia Bielefeld 19] 3 0 2 1 2 3 2 [20 20 Bo
russia Moenchengladbach 20] 3 0 2 1 1 3 2 [21 21 Schalke 21]
3 0 2 1 4 8 2 [22 22 Werder Bremen 22] 3 0 1 2 4 6 1 [23 23 MS
V Duisburg 23] 3 0 0 3 1 8 0

1: 11867 Germany
2: 3295 Bonn
3: 11867 Germany
4: 172326 FC_Bayern_Munich
5: 686142 Bundesliga
6: 497888 1._FC_Köln
7: 172326 FC_Bayern_Munich
8: 737608 VfB_Stuttgart
9: 331715 Borussia_Dortmund
10: 535879 Hamburger_SV
11: 502005 Bayer_04_Leverkusen
12: 513623 VfL_Bochum
13: 1734592 Karlsruher_SC
14: 881244 FC_St._Pauli
15: 1007441 TSV_1860_München
16: 1438835 SC_Freiburg
17: 1949688 Fortuna_Düsseldorf
18: 599978 F.C._Hansa_Rostock
19: 1145228 Arminia_Bielefeld
20: 263466 Borussia_Mönchengladbach
21: 410348 FC_Schalke_04
22: 7598082 SV_Werder_Bremen
23: 1586579 MSV_Duisburg

1: 11867 Germany
[pem_pos=1, final_scores=0.7868894,
final_local_scores=0.8658543, pem=-1.0465823,
similarity=35.62001]
2: 3295 Bonn
[pem_pos=0, final_scores=0.3964849,
final_local_scores=0.46297884, pem=0.0, similarity=17.905926]
3: 11867 Germany
[pem_pos=1, final_scores=0.71402687,
final_local_scores=0.78626317, pem=-1.0465823,
similarity=33.244854]
4: 172326 FC_Bayern_Munich
[pem_pos=0, final_scores=0.43585783,
final_local_scores=0.37351078, pem=0.0, similarity=15.23602]

5: 686142 Bundesliga
[pem_pos=0, final_scores=0.3689838,
final_local_scores=0.29910633, pem=-0.08715299,
similarity=13.489594]
6: 497888 1._FC_Köln
[pem_pos=10, final_scores=0.70684534,
final_local_scores=0.6675575, pem=-3.09805,
similarity=40.85864]
7: 172326 FC_Bayern_Munich
[pem_pos=0, final_scores=0.42682773,
final_local_scores=0.36364675, pem=0.0, similarity=14.941657]
8: 737608 VfB_Stuttgart
[pem_pos=0, final_scores=0.45055944,
final_local_scores=0.39068013, pem=0.0, similarity=15.748387]
9: 331715 Borussia_Dortmund
[pem_pos=0, final_scores=0.47370762,
final_local_scores=0.41318995, pem=0.0, similarity=16.420126]
10: 535879 Hamburger_SV
[pem_pos=1, final_scores=0.78628165,
final_local_scores=0.7526883, pem=-2.820485,
similarity=41.88967]
11: 502005 Bayer_04_Leverkusen
[pem_pos=0, final_scores=0.5179785,
final_local_scores=0.4619575, pem=0.0, similarity=17.875446]
12: 513623 VfL_Bochum
[pem_pos=0, final_scores=0.55663365,
final_local_scores=0.51337427, pem=0.0, similarity=19.409824]
13: 1734592 Karlsruher_SC
[pem_pos=1, final_scores=0.67825645,
final_local_scores=0.64314467, pem=-2.2263577,
similarity=35.389713]
14: 881244 FC_St._Pauli
[pem_pos=1, final_scores=0.5541174,
final_local_scores=0.51981974, pem=-0.11686432,
similarity=20.237698]
15: 1007441 TSV_1860_München
[pem_pos=0, final_scores=0.46654767,
final_local_scores=0.4149986, pem=0.0, similarity=16.474098]
16: 1438835 SC_Freiburg
[pem_pos=2, final_scores=0.82592326,
final_local_scores=0.822344, pem=-1.7450409,
similarity=38.1199]
17: 1949688 Fortuna_Düsseldorf
[pem_pos=0, final_scores=0.62940603,
final_local_scores=0.59299785, pem=0.0, similarity=21.785952]
18: 599978 F.C._Hansa_Rostock
[pem_pos=0, final_scores=0.5515744,
final_local_scores=0.5136526, pem=0.0, similarity=19.41813]
19: 1145228 Arminia_Bielefeld
[pem_pos=0, final_scores=0.51131237,
final_local_scores=0.4647237, pem=0.0, similarity=17.957994]
20: 263466 Borussia_Mönchengladbach

```
[pem_pos=0, final_scores=0.53096056,
final_local_scores=0.47653288, pem=0.0, similarity=18.310404]
21: 410348 FC_Schalke_04
[pem_pos=0, final_scores=0.5447096,
final_local_scores=0.4895091, pem=0.0, similarity=18.697641]
22: 7598082 SV_Werder_Bremen
[pem_pos=0, final_scores=0.61041206,
final_local_scores=0.5654225, pem=0.0, similarity=20.963047]
23: 1586579 MSV_Duisburg
[pem_pos=0, final_scores=0.5346969,
final_local_scores=0.49376625, pem=0.0, similarity=18.824682]
```

Sample: b'360_SOCCER'

```
SOCCER - [1 1 SWISS 1] PREMIER DIVISION RESULTS /
STANDINGS . [2 2 GENEVA 2] 1996-08-25 Results of [3 3 Swiss 3]
premier division matches played at the weekend : [4 4 Aarau 4]
1 [5 5 Young Boys 5] 0 [6 6 Grasshopper 6] 2 [7 7 Lucerne 7]
2 [8 8 Lugano 8] 1 Basle 1 Neuchatel 3 [9 9 St Gallen 9] 0 [10
10 Sion 10] 3 [11 11 Servette 11] 1 Standings ( tabulated
under played , won , drawn , lost , goals for , against ,
points ) : Neuchatel 8 6 1 1 12 7 19 [12 12 Grasshopper 12] 9
4 4 1 17 11 16 St. Gallen 9 4 4 1 6 5 16 [13 13 Lausanne 13]
9 4 2 3 18 13 14 [14 14 Aarau 14] 8 4 1 3 9 4 13 [15 15 Sion 1
5] 9 3 4 2 13 11 13 Zurich 9 2 5 2 9 9 11 Basle 8 2 3 3 12 11
9 [16 16 Servette 16] 9 2 3 4 10 12 9 [17 17 Lucerne 17] 8 1
5 2 10 11 8 [18 18 Lugano 18] 9 1 4 4 6 15 7 [19 19 Young Boys
19] 9 1 0 8 6 19 3
1: 26748 Switzerland
2: 12521 Geneva
3: 26748 Switzerland
4: 4663524 FC_Aarau
5: 861367 BSC_Young_Boys
6: 1573615 Grasshopper_Club_Zürich
7: 2841074 FC_Luzern
8: 5728757 FC_Lugano
9: 1607312 FC_St._Gallen
10: 1061075 FC_Sion
11: 893512 Servette_FC
12: 1573615 Grasshopper_Club_Zürich
13: 2092159 FC_Lausanne-Sport
14: 4663524 FC_Aarau
15: 1061075 FC_Sion
16: 893512 Servette_FC
17: 2841074 FC_Luzern
18: 5728757 FC_Lugano
19: 861367 BSC_Young_Boys
1: 26748 Switzerland
[pem_pos=0, final_scores=0.6542308,
final_local_scores=0.70928985, pem=-0.1392413,
similarity=26.013554]
2: 12521 Geneva
[pem_pos=0, final_scores=0.43116635,
```

```
final_local_scores=0.49922788, pem=-0.24772356,
similarity=20.334827]
3: 26748 Switzerland
    [pem_pos=0, final_scores=0.7847318,
final_local_scores=0.85184246, pem=-0.1392413,
similarity=30.267607]
4: 4663524 FC_Aarau
    [pem_pos=1, final_scores=0.6472521,
final_local_scores=0.62010896, pem=-0.831114,
similarity=27.114729]
5: 861367 BSC_Young_Boys
    [pem_pos=0, final_scores=0.5427764,
final_local_scores=0.49950176, pem=0.0, similarity=18.995842]
6: 1573615 Grasshopper_Club_Zürich
    [pem_pos=0, final_scores=0.61803085,
final_local_scores=0.5784143, pem=-1.0502282,
similarity=27.062054]
7: 2841074 FC_Luzern
    [pem_pos=6, final_scores=0.62799853,
final_local_scores=0.59536415, pem=-2.7928877,
similarity=37.044724]
8: 5728757 FC_Lugano
    [pem_pos=2, final_scores=0.3376596,
final_local_scores=0.28225192, pem=-1.772173,
similarity=22.150026]
9: 1607312 FC_St._Gallen
    [pem_pos=1, final_scores=0.35665256,
final_local_scores=0.31062806, pem=-0.54538625,
similarity=16.325369]
10: 1061075 FC_Sion
    [pem_pos=1, final_scores=0.8721879,
final_local_scores=0.8660323, pem=-1.4381809,
similarity=37.754898]
11: 893512 Servette_FC
    [pem_pos=0, final_scores=0.4579832,
final_local_scores=0.41354078, pem=0.0, similarity=16.430595]
12: 1573615 Grasshopper_Club_Zürich
    [pem_pos=0, final_scores=0.63422847,
final_local_scores=0.5961078, pem=-1.0502282,
similarity=27.590061]
13: 2092159 FC_Lausanne-Sport
    [pem_pos=3, final_scores=0.5401079,
final_local_scores=0.5001623, pem=-2.4908664,
similarity=32.56127]
14: 4663524 FC_Aarau
    [pem_pos=1, final_scores=0.6524674,
final_local_scores=0.625806, pem=-0.831114,
similarity=27.284739]
15: 1061075 FC_Sion
    [pem_pos=1, final_scores=0.85737866,
final_local_scores=0.8498555, pem=-1.4381809,
similarity=37.27215]
```

```
16: 893512 Servette_FC
    [pem_pos=0, final_scores=0.46578532,
final_local_scores=0.42206347, pem=0.0, similarity=16.684929]
17: 2841074 FC_Luzern
    [pem_pos=6, final_scores=0.64563936,
final_local_scores=0.6146341, pem=-2.7928877,
similarity=37.619778]
18: 5728757 FC_Lugano
    [pem_pos=2, final_scores=0.33555883,
final_local_scores=0.2799572, pem=-1.772173,
similarity=22.081547]
19: 861367 BSC_Young_Boys
    [pem_pos=0, final_scores=0.51224214,
final_local_scores=0.46614772, pem=0.0, similarity=18.00049]
```

Sample: b'361_GOLF'

GOLF - LEADING PRIZE MONEY WINNERS ON [1 1 EUROPEAN TOUR 1] .
[2 2 STUTTGART 2] , [3 3 Germany 3] 1996-08-25 Leading prize
money winners on the [4 4 European Tour 4] after Sunday 's [5
5 German Open 5] ([6 6 Britain 6] unless stated) : 1. [7 7 I
an Woosnam 7] 480,618 pounds sterling 2. [8 8 Colin Montgomer
ie 8] 429,449 3. [9 9 Lee Westwood 9] 301,972 4. [10 10 Robert
Allenby 10] ([11 11 Australia 11]) 291,088 5. [12 12 Mark Mc
Nulty 12] ([13 13 Zimbabwe 13]) 254,247 6. [14 14 Costantino
Rocca 14] ([15 15 Italy 15]) 253,337 7. [16 16 Andrew Colta
rt 16] 246,077 8. [17 17 Wayne Riley 17] ([18 18 Australia 18
]) 233,713 9. [19 19 Raymond Russell 19] 229,360 10. [20 20 S
tephen Ames 20] (Trinidad) 211,175 11. [21 21 Frank Nobilo 2
1] ([22 22 New Zealand 22]) 209,412 12. [23 23 Paul McGinley
23] ([24 24 Ireland 24]) 208,978 13. [25 25 Paul Lawrie 25]
207,990 14. [26 26 Padraig Harrington 26] ([27 27 Ireland 27]
) 202,593 15. [28 28 Retief Goosen 28] ([29 29 South Africa 2
9]) 188,143 16. [30 30 Jonathan Lomas 30] 181,005 17. [31 31
Paul Broadhurst 31] 172,580 18. [32 32 Peter Mitchell 32]
170,952 19. [33 33 Jim Payne 33] 165,150 20. [34 34 Russell Cl
aydon 34] 156,996

```
1: 841827 PGA_European_Tour
2: 28565 Stuttgart
3: 11867 Germany
4: 841827 PGA_European_Tour
5: 4488686 German_Open_(golf)
6: 31717 United_Kingdom
7: 1034991 Ian_Woosnam
8: 379616 Colin_Montgomerie
9: 667682 Lee_Westwood
10: 1983254 Robert_Allenby
11: 4689264 Australia
12: 1601826 Mark_McNulty
13: 34399 Zimbabwe
14: 2157151 Costantino_Rocca
15: 14532 Italy
16: 2150917 Andrew_Coltart
```

```
17: 5243170 Wayne_Riley
18: 4689264 Australia
19: 4375906 Raymond_Russell
20: 780871 Stephen_Ames
21: 2487707 Frank_Nobile
22: 4913064 New_Zealand
23: 1342124 Paul_McGinley
24: 14560 Republic_of_Ireland
25: 1564436 Paul_Lawrie
26: 1157761 Pádraig_Harrington
27: 14560 Republic_of_Ireland
28: 630525 Retief_Goosen
29: 17416221 South_Africa
30: 5272617 Jonathan_Lomas
31: 1688190 Paul_Broadhurst
32: 5000237 Peter_Mitchell_(golfer)
33: 4449258 Jim_Payne_(golfer)
34: 4291546 Russell_Claydon
1: 841827 PGA_European_Tour
    [pem_pos=0, final_scores=0.48313475,
final_local_scores=0.46027374, pem=0.0, similarity=17.8252]
2: 28565 Stuttgart
    [pem_pos=0, final_scores=0.49421686,
final_local_scores=0.58217156, pem=-0.15141529,
similarity=22.286293]
3: 11867 Germany
    [pem_pos=0, final_scores=0.7792155,
final_local_scores=0.8708913, pem=-0.15529114,
similarity=30.923344]
4: 841827 PGA_European_Tour
    [pem_pos=0, final_scores=0.49179983,
final_local_scores=0.46973902, pem=0.0, similarity=18.107662]
5: 4488686 German_Open_(golf)
    [pem_pos=0, final_scores=0.34240633,
final_local_scores=0.38270825, pem=-0.62093085,
similarity=18.887209]
6: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5331332,
final_local_scores=0.5972455, pem=-0.61114657,
similarity=25.23622]
7: 1034991 Ian_Woosnam
    [pem_pos=0, final_scores=0.75814,
final_local_scores=0.7891995, pem=0.0, similarity=27.640999]
8: 379616 Colin_Montgomerie
    [pem_pos=0, final_scores=0.81503445,
final_local_scores=0.84015363, pem=0.0, similarity=29.161573]
9: 667682 Lee_Westwood
    [pem_pos=0, final_scores=0.57388407,
final_local_scores=0.58881265, pem=0.0, similarity=21.661057]
10: 1983254 Robert_Allenby
    [pem_pos=0, final_scores=0.7833223,
final_local_scores=0.8390291, pem=0.0, similarity=29.128014]
```

11: 4689264 Australia
[pem_pos=0, final_scores=0.65320665,
final_local_scores=0.71867245, pem=-0.17168477,
similarity=26.469982]
12: 1601826 Mark_McNulty
[pem_pos=0, final_scores=0.66558623,
final_local_scores=0.7081169, pem=0.0, similarity=25.221333]
13: 34399 Zimbabwe
[pem_pos=0, final_scores=0.49965626,
final_local_scores=0.58293736, pem=-0.36737105,
similarity=23.483547]
14: 2157151 Costantino_Rocca
[pem_pos=0, final_scores=0.7790703,
final_local_scores=0.83677846, pem=0.0, similarity=29.060852]
15: 14532 Italy
[pem_pos=0, final_scores=0.66634387,
final_local_scores=0.76260173, pem=-0.20612453,
similarity=27.968208]
16: 2150917 Andrew_Coltart
[pem_pos=0, final_scores=0.7214669,
final_local_scores=0.75845087, pem=0.0, similarity=26.7234]
17: 5243170 Wayne_Riley
[pem_pos=0, final_scores=0.63488233,
final_local_scores=0.6858653, pem=0.0, similarity=24.5573]
18: 4689264 Australia
[pem_pos=0, final_scores=0.6307976,
final_local_scores=0.694194, pem=-0.17168477,
similarity=25.739496]
19: 4375906 Raymond_Russell
[pem_pos=0, final_scores=0.7272429,
final_local_scores=0.7975206, pem=0.0, similarity=27.889318]
20: 780871 Stephen_Ames
[pem_pos=0, final_scores=0.6276312,
final_local_scores=0.66847557, pem=0.0, similarity=24.038357]
21: 2487707 Frank_Nobilo
[pem_pos=0, final_scores=0.7677396,
final_local_scores=0.8157351, pem=0.0, similarity=28.432875]
22: 4913064 New_Zealand
[pem_pos=0, final_scores=0.58463484,
final_local_scores=0.65058315, pem=-0.23795462,
similarity=24.798447]
23: 1342124 Paul_McGinley
[pem_pos=0, final_scores=0.63357705,
final_local_scores=0.65573746, pem=0.0, similarity=23.658226]
24: 14560 Republic_of_Ireland
[pem_pos=1, final_scores=0.6044587,
final_local_scores=0.6713852, pem=-1.5040284,
similarity=32.30433]
25: 1564436 Paul_Lawrie
[pem_pos=0, final_scores=0.71731776,
final_local_scores=0.7625605, pem=0.0, similarity=26.846039]
26: 1157761 Pádraig_Harrington

```
[pem_pos=0, final_scores=0.5794322,
final_local_scores=0.6001598, pem=0.0, similarity=21.99968]
27: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.5440368,
final_local_scores=0.60538346, pem=-1.5040284,
similarity=30.334703]
28: 630525 Retief_Goosen
    [pem_pos=0, final_scores=0.6420015,
final_local_scores=0.6884992, pem=0.0, similarity=24.635902]
29: 17416221 South_Africa
    [pem_pos=0, final_scores=0.5822632,
final_local_scores=0.65409034, pem=-0.26251382,
similarity=25.036665]
30: 5272617 Jonathan_Lomas
    [pem_pos=0, final_scores=0.6111402,
final_local_scores=0.6865257, pem=0.0, similarity=24.57701]
31: 1688190 Paul_Broadhurst
    [pem_pos=0, final_scores=0.7653032,
final_local_scores=0.830379, pem=0.0, similarity=28.869879]
32: 5000237 Peter_Mitchell_(golfer)
    [pem_pos=3, final_scores=0.49508375,
final_local_scores=0.5516764, pem=-1.2033675,
similarity=27.096937]
33: 4449258 Jim_Payne_(golfer)
    [pem_pos=0, final_scores=0.6635477,
final_local_scores=0.7394908, pem=0.0, similarity=26.157593]
34: 4291546 Russell_Claydon
    [pem_pos=0, final_scores=0.6514161,
final_local_scores=0.7194586, pem=0.0, similarity=25.559792]
```

Sample: b'362_CYCLING'

```
CYCLING - [1 1 SWISS 1] GRAND PRIX RESULT . [2 2 ZURICH 2]
1996-08-25 Leading results in the 232-km [3 3 Swiss 3] Grand
Prix 1 World Cup cycling race on Sunday : 1. [4 4 Andrea Ferr
igato 4] ( [5 5 Italy 5] ) 5 hours 51 minutes 52 seconds 2. [6
6 Michele Bartoli 6] ( [7 7 Italy 7] ) 3. [8 8 Johan Museeuw 8
] ( [9 9 Belgium 9] ) 4. [10 10 Lance Armstrong 10] ( [11 11 U
.S. 11] ) 5. [12 12 Francesco Casagrande 12] ( [13 13 Italy 13
] ) 6. Alessandro Baronti ( [14 14 Italy 14] ) 7. [15 15 Frank
Vandenbroucke 15] ( [16 16 Belgium 16] ) all same time 8. [17
17 Fabio Baldato 17] ( [18 18 Italy 18] ) 11 seconds behind
9. [19 19 Maurizio Fondriest 19] ( [20 20 Italy 20] ) 10. [21
21 Laurent Jalabert 21] ( [22 22 France 22] ) both same time
Leading 2 World Cup standings ( after 8 of 11 rounds ) : 1. [2
3 23 Museeuw 23] 162 points 2. [24 24 Ferrigato 24] 112 3. [25
25 Bartoli 25] 108 4. [26 26 Stefano Zanini 26] ( [27 27 Italy
27] ) 88 5. [28 28 Armstrong 28] 81 6. [29 29 Baldato 29] 77
7. [30 30 Alexandre Gontchenkov 30] ( [31 31 Ukraine 31] ) 67
8. [32 32 Gabriele Colombo 32] ( [33 33 Italy 33] ) 58 9. [34
34 Andrei Tchmil 34] ( [35 35 Ukraine 35] ) 56 10. [36 36 Max
Sciandri 36] ( [37 37 Britain 37] ) 55
1: 26748 Switzerland
```

2: 40334603 Zürich
3: 26748 Switzerland
4: 14056108 Andrea_Ferrigato
5: 14532 Italy
6: 2912926 Michele_Bartoli
7: 14532 Italy
8: 859159 Johan_Museeuw
9: 3343 Belgium
10: 23243880 Lance_Armstrong
11: 3434750 United_States
12: 2204856 Francesco_Casagrande
13: 14532 Italy
14: 14532 Italy
15: 2399432 Frank_Vandenbroucke_(cyclist)
16: 3343 Belgium
17: 12841333 Fabio_Baldato
18: 14532 Italy
19: 4010271 Maurizio_Fondriest
20: 14532 Italy
21: 909249 Laurent_Jalabert
22: 5843419 France
23: 859159 Johan_Museeuw
24: 14056108 Andrea_Ferrigato
25: 2912926 Michele_Bartoli
26: 6232798 Stefano_Zanini
27: 14532 Italy
28: 23243880 Lance_Armstrong
29: 12841333 Fabio_Baldato
30: 17143953 Alexander_Gontchenkov
31: 31750 Ukraine
32: 15464425 Gabriele_Colombo
33: 14532 Italy
34: 2547354 Andrei_Tchmil
35: 31750 Ukraine
36: 11438071 Max_Sciandri
37: 31717 United_Kingdom
1: 2036254 UCI_Road_World_Cup
 [winner: 41395970 Cup, pem_pos=0,
final_scores=-0.6837118, final_local_scores=-0.5020281,
pem=-1.2544911, similarity=-4.069689]
 RECALL MISS: 2036254 UCI_Road_World_Cup
 [winner: 11242642 World_(Bee_Gees_song), pem_pos=6,
final_scores=-1.7607465, final_local_scores=-1.7590435,
pem=-3.2688804, similarity=-30.626972]
 RECALL MISS: 2036254 UCI_Road_World_Cup]
2: 2036254 UCI_Road_World_Cup
 [winner: 11370 FIFA_World_Cup, pem_pos=0,
final_scores=-0.24057256, final_local_scores=-0.25409,
pem=-1.001087, similarity=1.9512272]
 RECALL MISS: 2036254 UCI_Road_World_Cup]
1: 26748 Switzerland
 [pem_pos=0, final_scores=0.5054663,

```
final_local_scores=0.4989205, pem=-0.1392413,
similarity=19.735712]
2: 40334603 Zürich
    [pem_pos=0, final_scores=0.6011309,
final_local_scores=0.6223604, pem=-0.08839822,
similarity=23.142912]
3: 26748 Switzerland
    [pem_pos=0, final_scores=0.5156248,
final_local_scores=0.5100172, pem=-0.1392413,
similarity=20.06686]
4: 14056108 Andrea_Ferrigato
    [pem_pos=0, final_scores=0.61498064,
final_local_scores=0.731906, pem=0.0, similarity=25.931248]
5: 14532 Italy
    [pem_pos=0, final_scores=0.7772384,
final_local_scores=0.78029853, pem=-0.20612453,
similarity=28.496315]
6: 2912926 Michele_Bartoli
    [pem_pos=0, final_scores=0.9181306,
final_local_scores=1.0153875, pem=0.0, similarity=34.390903]
7: 14532 Italy
    [pem_pos=0, final_scores=0.7839727,
final_local_scores=0.7876547, pem=-0.20612453,
similarity=28.715837]
8: 859159 Johan_Museeuw
    [pem_pos=0, final_scores=0.7068165,
final_local_scores=0.78007716, pem=0.0, similarity=27.36877]
9: 3343 Belgium
    [pem_pos=0, final_scores=0.6983884,
final_local_scores=0.70637196, pem=-0.24180195,
similarity=26.484219]
10: 23243880 Lance_Armstrong
    [pem_pos=0, final_scores=0.62241536,
final_local_scores=0.67098397, pem=0.0, similarity=24.113213]
11: 3434750 United_States
    [pem_pos=0, final_scores=0.44167596,
final_local_scores=0.47495955, pem=-0.14075686,
similarity=19.02891]
12: 2204856 Francesco_Casagrande
    [pem_pos=0, final_scores=0.56317496,
final_local_scores=0.66113895, pem=0.0, similarity=23.819418]
13: 14532 Italy
    [pem_pos=0, final_scores=0.79690844,
final_local_scores=0.80178505, pem=-0.20612453,
similarity=29.137516]
14: 14532 Italy
    [pem_pos=0, final_scores=0.7949535,
final_local_scores=0.7996496, pem=-0.20612453,
similarity=29.07379]
15: 2399432 Frank_Vandenbroucke_(cyclist)
    [pem_pos=0, final_scores=0.6899243,
final_local_scores=0.7781617, pem=0.0, similarity=27.311611]
```

16: 3343 Belgium
[pem_pos=0, final_scores=0.6997058,
final_local_scores=0.70781094, pem=-0.24180195,
similarity=26.52716]
17: 12841333 Fabio_Baldato
[pem_pos=0, final_scores=0.7163948,
final_local_scores=0.82470036, pem=0.0, similarity=28.700417]
18: 14532 Italy
[pem_pos=0, final_scores=0.7775714,
final_local_scores=0.78066224, pem=-0.20612453,
similarity=28.50717]
19: 4010271 Maurizio_Fondriest
[pem_pos=0, final_scores=0.63979304,
final_local_scores=0.7423515, pem=0.0, similarity=26.242962]
20: 14532 Italy
[pem_pos=0, final_scores=0.79693705,
final_local_scores=0.80181634, pem=-0.20612453,
similarity=29.13845]
21: 909249 Laurent_Jalabert
[pem_pos=0, final_scores=0.89984566,
final_local_scores=0.984425, pem=0.0, similarity=33.46692]
22: 5843419 France
[pem_pos=0, final_scores=0.67432123,
final_local_scores=0.6877772, pem=-0.16240834,
similarity=25.497559]
23: 859159 Johan_Museeuw
[pem_pos=0, final_scores=0.6520691,
final_local_scores=0.7202739, pem=0.0, similarity=25.584124]
24: 14056108 Andrea_Ferrigato
[pem_pos=0, final_scores=0.5098017,
final_local_scores=0.617014, pem=0.0, similarity=22.502642]
25: 2912926 Michele_Bartoli
[pem_pos=0, final_scores=0.74041027,
final_local_scores=0.8301321, pem=0.0, similarity=28.86251]
26: 6232798 Stefano_Zanini
[pem_pos=0, final_scores=0.5965442,
final_local_scores=0.7041025, pem=0.0, similarity=25.101536]
27: 14532 Italy
[pem_pos=0, final_scores=0.79689914,
final_local_scores=0.8017749, pem=-0.20612453,
similarity=29.137213]
28: 23243880 Lance_Armstrong
[pem_pos=0, final_scores=1.1996332,
final_local_scores=1.3093169, pem=0.0, similarity=43.162342]
29: 12841333 Fabio_Baldato
[pem_pos=0, final_scores=0.69868624,
final_local_scores=0.80535644, pem=0.0, similarity=28.123156]
30: 17143953 Alexander_Gontchenkov
[pem_pos=0, final_scores=0.63203925,
final_local_scores=0.7787771, pem=0.0, similarity=27.329975]
31: 31750 Ukraine
[pem_pos=0, final_scores=0.6085885,

```
final_local_scores=0.6553294, pem=-0.35407752,
similarity=25.571577]
32: 15464425 Gabriele_Colombo
    [pem_pos=0, final_scores=0.42081356,
final_local_scores=0.5705372, pem=0.0, similarity=21.11568]
33: 14532 Italy
    [pem_pos=0, final_scores=0.75601846,
final_local_scores=0.757119, pem=-0.20612453,
similarity=27.804592]
34: 2547354 Andrei_Tchmil
    [pem_pos=0, final_scores=0.8399063,
final_local_scores=0.95202094, pem=0.0, similarity=32.499916]
35: 31750 Ukraine
    [pem_pos=0, final_scores=0.5612975,
final_local_scores=0.60367113, pem=-0.35407752,
similarity=24.029991]
36: 11438071 Max_Sciandri
    [pem_pos=0, final_scores=0.80871016,
final_local_scores=0.93105334, pem=0.0, similarity=31.8742]
37: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.485138,
final_local_scores=0.52815515, pem=-0.61114657,
similarity=23.174425]
```

Sample: b'363_CYCLING'

CYCLING - [1 1 FERRIGATO 1] SPRINTS TO SECOND STRAIGHT WORLD CUP WIN . [2 2 ZURICH 2] 1996-08-25 [3 3 Andrea Ferrigato 3] of [4 4 Italy 4] sprinted to his second cycling 1 World Cup win in successive weekends with victory in the [5 5 Swiss 5] Grand Prix on Sunday . [6 6 Ferrigato 6] , winner of the [7 7 Leeds Classic 7] last Sunday with a one second win over [8 8 Britain 8] 's [9 9 Max Sciandri 9] , posted a similarly narrow margin of victory again . The 26-year-old [10 10 Italian 10] surged past compatriot [11 11 Michele Bartoli 11] and last year 's winner and defending 2 World Cup champion [12 12 Johan Museeuw 12] of [13 13 Belgium 13] in the final few metres of the 237km race . All three clocked the same time of five hours 51 minutes , 52 seconds . Former world champion [14 14 Lance Armstrong 14] of the [15 15 United States 15] was in front as the leading pack of seven riders turned into the [16 16 Oerlikon 16] velodrome for the final one lap sprint but quickly faded and settled for fourth . The back-to-back wins vault [17 17 Ferrigato 17] from sixth to second in the overall 3 World Cup rankings with 112 points but [18 18 Museeuw 18] continues to hold a commanding lead with 162 points after eight of the 11 rounds .

```
1: 14056108 Andrea_Ferrigato
2: 40334603 Zürich
3: 14056108 Andrea_Ferrigato
4: 14532 Italy
5: 26748 Switzerland
6: 14056108 Andrea_Ferrigato
```

```
7: 14016488 Wincanton_Classic
8: 31717 United_Kingdom
9: 11438071 Max_Sciandri
10: 14532 Italy
11: 2912926 Michele_Bartoli
12: 859159 Johan_Museeuw
13: 3343 Belgium
14: 23243880 Lance_Armstrong
15: 3434750 United_States
16: 3369900 Oerlikon_(Zürich)
17: 14056108 Andrea_Ferrigato
18: 859159 Johan_Museeuw
1: 2036254 UCI_Road_World_Cup
    [winner: 11370 FIFA_World_Cup, pem_pos=0,
final_scores=-0.1640922, final_local_scores=-0.15343097,
pem=-1.001087, similarity=4.9550934]
    RECALL MISS: 2036254 UCI_Road_World_Cup]
2: 2036254 UCI_Road_World_Cup
    [winner: 11370 FIFA_World_Cup, pem_pos=0,
final_scores=-0.11213182, final_local_scores=-0.09667215,
pem=-1.001087, similarity=6.64889]
    RECALL MISS: 2036254 UCI_Road_World_Cup]
3: 2036254 UCI_Road_World_Cup
    [winner: 183628 FIS_Alpine_Ski_World_Cup, pem_pos=2,
final_scores=-0.22013953, final_local_scores=-0.24136993,
pem=-2.7628422, similarity=11.911519]
    RECALL MISS: 2036254 UCI_Road_World_Cup]
1: 14056108 Andrea_Ferrigato
    [pem_pos=0, final_scores=0.5208654,
final_local_scores=0.5857591, pem=0.0, similarity=21.569933]
2: 40334603 Zürich
    [pem_pos=0, final_scores=0.6375135,
final_local_scores=0.66723156, pem=-0.08839822,
similarity=24.481956]
3: 14056108 Andrea_Ferrigato
    [pem_pos=0, final_scores=0.6977133,
final_local_scores=0.7789386, pem=0.0, similarity=27.334795]
4: 14532 Italy
    [pem_pos=0, final_scores=0.75007296,
final_local_scores=0.78418326, pem=-0.20612453,
similarity=28.612244]
5: 26748 Switzerland
    [pem_pos=0, final_scores=0.62473243,
final_local_scores=0.6447192, pem=-0.1392413,
similarity=24.086636]
6: 14056108 Andrea_Ferrigato
    [pem_pos=0, final_scores=0.4993052,
final_local_scores=0.5622079, pem=0.0, similarity=20.867117]
7: 14016488 Wincanton_Classic
    [pem_pos=0, final_scores=0.33402342,
final_local_scores=0.49298877, pem=0.0, similarity=18.801481]
8: 31717 United_Kingdom
```

```

[pem_pos=0, final_scores=0.5830894,
final_local_scores=0.65103656, pem=-0.61114657,
similarity=26.841454]
9: 11438071 Max_Sciandri
    [pem_pos=0, final_scores=0.6215516,
final_local_scores=0.7217771, pem=0.0, similarity=25.62898]
10: 14532 Italy
    [pem_pos=0, final_scores=0.59381175,
final_local_scores=0.61349165, pem=-0.75948656,
similarity=26.527733]
11: 2912926 Michele_Bartoli
    [pem_pos=0, final_scores=0.53715414,
final_local_scores=0.62692446, pem=0.0, similarity=22.79839]
12: 859159 Johan_Museeuw
    [pem_pos=0, final_scores=0.56359,
final_local_scores=0.6311603, pem=0.0, similarity=22.924795]
13: 3343 Belgium
    [pem_pos=0, final_scores=0.6244361,
final_local_scores=0.6466241, pem=-0.24180195,
similarity=24.701223]
14: 23243880 Lance_Armstrong
    [pem_pos=0, final_scores=0.52295035,
final_local_scores=0.58423615, pem=0.0, similarity=21.524485]
15: 3434750 United_States
    [pem_pos=0, final_scores=0.48562145,
final_local_scores=0.54084855, pem=-0.073457584,
similarity=20.629187]
16: 3369900 Oerlikon_(Zürich)
    [pem_pos=2, final_scores=0.24298967,
final_local_scores=0.37657768, pem=-1.1003703,
similarity=21.311525]
17: 14056108 Andrea_Ferrigato
    [pem_pos=0, final_scores=0.59653753,
final_local_scores=0.6684194, pem=0.0, similarity=24.036682]
18: 859159 Johan_Museeuw
    [pem_pos=0, final_scores=0.8047696,
final_local_scores=0.8946124, pem=0.0, similarity=30.78673]

```

Sample: b'364_GOLF'

GOLF - [1 1 GERMAN OPEN 1] SCORES . [2 2 STUTTGART 2] , [3 3 Germany 3] 1996-08-25 [4 4 Briton 4] [5 5 Ian Woosnam 5] won the [6 6 German Open 6] golf championship on Sunday after the final round was abandoned because of torrential rain . Scores after three rounds ([7 7 Britain 7] unless stated) : 193 [8 8 Ian Woosnam 8] 64 64 65 . 199 Thomas Gogele ([9 9 Germany 9]) 67 65 67 , [10 10 Robert Karlsson 10] ([11 11 Sweden 11]) 67 62 70 , [12 12 Ian Pyman 12] 66 64 69 , Fernando Roca ([13 13 Spain 13]) 66 64 69 . 200 [14 14 Diego Borrego 14] ([15 15 Spain 15]) 69 63 68 , [16 16 Miguel Angel Martin 16] ([17 17 Spain 17]) 66 66 68 . 201 [18 18 Stephen Ames 18] (Trinidad) 68 65 68 , [19 19 Roger Chapman 19] 72 62 67 , [20 20 Paul Broadhurst 20] 62 70 69 , Stephen Field 66 65 70 , [

21 21 Carl Suneson 21] ([22 22 Spain 22]) 65 66 70 202 [23 2
3 Greg Turner 23] ([24 24 New Zealand 24]) 70 67 65 , Heinz-
Peter Thul ([25 25 Germany 25]) 70 67 65 , [26 26 Ronan Raf-
ferty 26] 64 72 66 , [27 27 Barry Lane 27] 68 67 67 , [28 28 D-
avid Carter 28] 66 69 67 , [29 29 Michael Jonzon 29] ([30 30
Sweden 30]) 67 67 68 , David Williams 67 67 68 203 [31 31 Lee
Westwood 31] 66 71 66 , [32 32 Gary Emerson 32] 68 69 66 , [33
33 Peter Baker 33] 70 66 67 , [34 34 Des Smyth 34] ([35 35 Ir-
eland 35]) 66 69 68 , [36 36 Paul Lawrie 36] 66 69 68 ,
Francisco Cea ([37 37 Spain 37]) 68 66 69 , [38 38 Pedro Li-
nhart 38] ([39 39 Spain 39]) 67 67 69 , [40 40 Jonathan Lom-
as 40] 67 67 69 , [41 41 Paul Eales 41] 67 68 68 , [42 42 Ray-
mond Russell 42] 63 69 71
1: 4488686 German_Open_(golf)
2: 28565 Stuttgart
3: 11867 Germany
4: 19097669 British_people
5: 1034991 Ian_Woosnam
6: 4488686 German_Open_(golf)
7: 31717 United_Kingdom
8: 1034991 Ian_Woosnam
9: 11867 Germany
10: 1988802 Robert_Karlsson
11: 5058739 Sweden
12: 5088382 Iain_Pyman
13: 26667 Spain
14: 5271765 Diego_Borrego
15: 26667 Spain
16: 5022198 Miguel_Ángel_Martín_(golfer)
17: 26667 Spain
18: 780871 Stephen_Ames
19: 4380250 Roger_Chapman_(golfer)
20: 1688190 Paul_Broadhurst
21: 11818919 Carl_Suneson
22: 26667 Spain
23: 5276824 Greg_Turner
24: 4913064 New_Zealand
25: 11867 Germany
26: 1519060 Ronan_Rafferty
27: 2167839 Barry_Lane
28: 4344819 David_Carter_(golfer)
29: 5242970 Michael_Jonzon
30: 5058739 Sweden
31: 667682 Lee_Westwood
32: 5412184 Gary_Emerson
33: 4308246 Peter_Baker_(golfer)
34: 5193762 Des_Smyth
35: 14560 Republic_of_Ireland
36: 1564436 Paul_Lawrie
37: 26667 Spain
38: 5285076 Pedro_Linhart
39: 26667 Spain

```
40: 5272617 Jonathan_Lomas
41: 4435993 Paul_Eales
42: 4375906 Raymond_Russell
1: 4488686 German_Open_(golf)
    [pem_pos=0, final_scores=0.61955124,
final_local_scores=0.69053024, pem=-0.62093085,
similarity=28.07323]
2: 28565 Stuttgart
    [pem_pos=0, final_scores=0.5140263,
final_local_scores=0.57814956, pem=-0.15141529,
similarity=22.16627]
3: 11867 Germany
    [pem_pos=0, final_scores=0.80820924,
final_local_scores=0.8701572, pem=-0.15529114,
similarity=30.901438]
4: 19097669 British_people
    [pem_pos=0, final_scores=0.20901118,
final_local_scores=0.25163293, pem=-0.89919627,
similarity=16.48891]
5: 1034991 Ian_Woosnam
    [pem_pos=0, final_scores=0.65711004,
final_local_scores=0.6900547, pem=0.0, similarity=24.682322]
6: 4488686 German_Open_(golf)
    [pem_pos=0, final_scores=0.39269918,
final_local_scores=0.44272882, pem=-0.62093085,
similarity=20.67834]
7: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.4378336,
final_local_scores=0.4736995, pem=-0.61114657,
similarity=21.54936]
8: 1034991 Ian_Woosnam
    [pem_pos=0, final_scores=0.63474643,
final_local_scores=0.6656259, pem=0.0, similarity=23.953316]
9: 11867 Germany
    [pem_pos=0, final_scores=0.8145365,
final_local_scores=0.8770687, pem=-0.15529114,
similarity=31.107689]
10: 1988802 Robert_Karlsson
    [pem_pos=0, final_scores=0.7669783,
final_local_scores=0.8538538, pem=0.0, similarity=29.570414]
11: 5058739 Sweden
    [pem_pos=0, final_scores=0.7447464,
final_local_scores=0.78774005, pem=-0.2135508,
similarity=28.75877]
12: 5088382 Iain_Pyman
    [pem_pos=0, final_scores=0.8074433,
final_local_scores=0.9010026, pem=0.0, similarity=30.977427]
13: 26667 Spain
    [pem_pos=0, final_scores=0.57963234,
final_local_scores=0.6036583, pem=-0.18857051,
similarity=23.129557]
14: 5271765 Diego_Borrego
```

```
[pem_pos=0, final_scores=0.5346504,
final_local_scores=0.6210195, pem=0.0, similarity=22.622173]
15: 26667 Spain
    [pem_pos=0, final_scores=0.618195,
final_local_scores=0.64578223, pem=-0.18857051,
similarity=24.386618]
16: 5022198 Miguel_Ángel_Martín_(golfer)
    [pem_pos=0, final_scores=0.60409594,
final_local_scores=0.66959566, pem=0.0, similarity=24.071783]
17: 26667 Spain
    [pem_pos=0, final_scores=0.5808199,
final_local_scores=0.6049557, pem=-0.18857051,
similarity=23.168272]
18: 780871 Stephen_Ames
    [pem_pos=0, final_scores=0.63907313,
final_local_scores=0.69901025, pem=0.0, similarity=24.949574]
19: 4380250 Roger_Chapman_(golfer)
    [pem_pos=1, final_scores=0.69152635,
final_local_scores=0.7559486, pem=-0.5573004,
similarity=29.679415]
20: 1688190 Paul_Broadhurst
    [pem_pos=0, final_scores=0.79381067,
final_local_scores=0.8735402, pem=0.0, similarity=30.157894]
21: 11818919 Carl_Suneson
    [pem_pos=0, final_scores=0.7816779,
final_local_scores=0.85876656, pem=0.0, similarity=29.71702]
22: 26667 Spain
    [pem_pos=0, final_scores=0.6238951,
final_local_scores=0.6520088, pem=-0.18857051,
similarity=24.572432]
23: 5276824 Greg_Turner
    [pem_pos=0, final_scores=0.48080498,
final_local_scores=0.54991996, pem=0.0, similarity=20.500422]
24: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.5094522,
final_local_scores=0.56890225, pem=-0.23795462,
similarity=22.360926]
25: 11867 Germany
    [pem_pos=0, final_scores=0.7680858,
final_local_scores=0.82632834, pem=-0.15529114,
similarity=29.593496]
26: 1519060 Ronan_Rafferty
    [pem_pos=0, final_scores=0.45724016,
final_local_scores=0.49808168, pem=0.0, similarity=18.953465]
27: 2167839 Barry_Lane
    [pem_pos=0, final_scores=0.60572624,
final_local_scores=0.6701442, pem=0.0, similarity=24.088152]
28: 4344819 David_Carter_(golfer)
    [pem_pos=4, final_scores=0.6285978,
final_local_scores=0.686678, pem=-1.8641758,
similarity=34.71923]
29: 5242970 Michael_Jonzon
```

```
[pem_pos=0, final_scores=0.67201406,
final_local_scores=0.7606963, pem=0.0, similarity=26.790407]
30: 5058739 Sweden
    [pem_pos=0, final_scores=0.73781425,
final_local_scores=0.78016776, pem=-0.2135508,
similarity=28.532797]
31: 667682 Lee_Westwood
    [pem_pos=0, final_scores=0.58723736,
final_local_scores=0.62516546, pem=0.0, similarity=22.745897]
32: 5412184 Gary_Emerson
    [pem_pos=0, final_scores=0.48088038,
final_local_scores=0.55891997, pem=0.0, similarity=20.769]
33: 4308246 Peter_Baker_(golfer)
    [pem_pos=0, final_scores=0.4359458,
final_local_scores=0.48161125, pem=-0.44704348,
similarity=20.893045]
34: 5193762 Des_Smyth
    [pem_pos=0, final_scores=0.73330593,
final_local_scores=0.80844295, pem=0.0, similarity=28.215263]
35: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.5938457,
final_local_scores=0.6524231, pem=-1.5040284,
similarity=31.738459]
36: 1564436 Paul_Lawrie
    [pem_pos=0, final_scores=0.76712435,
final_local_scores=0.84245414, pem=0.0, similarity=29.230225]
37: 26667 Spain
    [pem_pos=0, final_scores=0.6246447,
final_local_scores=0.6528275, pem=-0.18857051,
similarity=24.596863]
38: 5285076 Pedro_Linhart
    [pem_pos=0, final_scores=0.6617206,
final_local_scores=0.744743, pem=0.0, similarity=26.31433]
39: 26667 Spain
    [pem_pos=0, final_scores=0.5968906,
final_local_scores=0.6225105, pem=-0.18857051,
similarity=23.692142]
40: 5272617 Jonathan_Lomas
    [pem_pos=0, final_scores=0.73626715,
final_local_scores=0.8206269, pem=0.0, similarity=28.578857]
41: 4435993 Paul_Eales
    [pem_pos=0, final_scores=0.67302763,
final_local_scores=0.7432267, pem=0.0, similarity=26.269081]
42: 4375906 Raymond_Russell
    [pem_pos=0, final_scores=0.65122515,
final_local_scores=0.7116889, pem=0.0, similarity=25.327929]
```

Sample: b'365_SOCCER'
SOCCER - [1 1 PSV 1] BEAT [2 2 GRONINGEN 2] 4-1 TO PULL AWAY
FROM 1 AJAX . [3 3 AMSTERDAM 3] 1996-08-25 [4 4 Belgian 4]
international [5 5 Luc Nilis 5] scored twice on Sunday as [6 6
PSV Eindhoven 6] came from behind to beat [7 7 Groningen 7]

4-1 in [8 8 Eindhoven 8] . [9 9 PSV 9] and [10 10 Vitesse Arnhem 10] are the only unbeaten teams after two rounds of the 2 Dutch league . Defending champions [11 11 Ajax Amsterdam 11] were defeated 2-0 loss away to [12 12 Heerenveen 12] on Saturday . [13 13 Groningen 13] took the lead in the seventh minute when [14 14 Dean Gorre 14] intercepted a back pass from [15 15 Ernest Faber 15] to goalkeeper [16 16 Ronald Waterreus 16] and shot home . [17 17 Faber 17] made amends in the 32nd minute when he headed in a corner to score the equaliser . [18 18 PSV 18] took control in the second half but could not score until [19 19 Groningen 19] striker Romano Sion was sent off in the 58th minute . Five minutes after his dismissal , [20 20 Nilis 20] gave [21 21 PSV 21] the lead and in the final 15 minutes he added another as did [22 22 Željko Petrović 22] . [23 23 Vitesse Arnhem 23] upstaged [24 24 Utrecht 24] 1-0 despite ending the match with only nine men following the dismissal of defenders [25 25 Raymond Atteveld 25] and Erwin van der Looi . [26 26 Gaston Taument 26] scored twice and newly signed [27 27 Argentine 27] [28 28 Pablo Sánchez 28] once in [29 29 Feyenoord Rotterdam 29] 's 3-0 victory over [30 30 Volendam 30] .

1: 79820 PSV_Eindhoven
2: 1103112 FC_Groningen
3: 844 Amsterdam
4: 3343 Belgium
5: 2945257 Luc_Nilis
6: 79820 PSV_Eindhoven
7: 1103112 FC_Groningen
8: 9420 Eindhoven
9: 79820 PSV_Eindhoven
10: 834256 Vitesse
11: 2273 AFC_Ajax
12: 745757 SC_Heerenveen
13: 1103112 FC_Groningen
14: 7632263 Dean_Gorré
15: 9426084 Ernest_Faber
16: 1456550 Ronald_Waterreus
17: 9426084 Ernest_Faber
18: 79820 PSV_Eindhoven
19: 1103112 FC_Groningen
20: 2945257 Luc_Nilis
21: 79820 PSV_Eindhoven
22: 2212066 Željko_Petrović
23: 834256 Vitesse
24: 834226 FC_Utrecht
25: 4287016 Raymond_Atteveld
26: 5824381 Gaston_Taument
27: 18951905 Argentine
28: 9420663 Pablo_Sánchez
29: 47707 Feyenoord
30: 2202202 FC_Volendam
1: 2273 AFC_Ajax

```
[winner: 1610950 Ajax_(programming), pem_pos=0,
final_scores=-0.38691136, final_local_scores=-0.2696588,
pem=0.0, similarity=-3.957444]
    RECALL MISS: 2273 AFC_Ajax]
2: 1332105 Eredivisie
    [winner: 9647657 Netherlands_national_football_team,
pem_pos=11, final_scores=-0.13055287,
final_local_scores=-0.21824479, pem=-4.675365,
similarity=23.002214]
    RECALL MISS: 1332105 Eredivisie]
1: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.7513955,
final_local_scores=0.7141841, pem=0.0, similarity=25.402391]
2: 1103112 FC_Groningen
    [pem_pos=2, final_scores=0.6900674,
final_local_scores=0.6559389, pem=-1.9677447,
similarity=34.365143]
3: 844 Amsterdam
    [pem_pos=0, final_scores=0.5410422,
final_local_scores=0.5934792, pem=-0.0642803,
similarity=22.149883]
4: 3343 Belgium
    [pem_pos=0, final_scores=0.6282157,
final_local_scores=0.70403224, pem=-0.12403572,
similarity=25.773964]
5: 2945257 Luc_Nilis
    [pem_pos=0, final_scores=0.8147376,
final_local_scores=0.88928586, pem=0.0, similarity=30.627777]
6: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.5296667,
final_local_scores=0.47197908, pem=0.0, similarity=18.174511]
7: 1103112 FC_Groningen
    [pem_pos=2, final_scores=0.8238923,
final_local_scores=0.8021223, pem=-1.9677447,
similarity=38.727543]
8: 9420 Eindhoven
    [pem_pos=0, final_scores=0.48599666,
final_local_scores=0.5215734, pem=0.0, similarity=19.654505]
9: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.5450186,
final_local_scores=0.48874873, pem=0.0, similarity=18.67495]
10: 834256 Vitesse
    [pem_pos=0, final_scores=0.6713977,
final_local_scores=0.64495826, pem=0.0, similarity=23.336554]
11: 2273 AFC_Ajax
    [pem_pos=0, final_scores=0.70604616,
final_local_scores=0.6709478, pem=0.0, similarity=24.112133]
12: 745757 SC_Heerenveen
    [pem_pos=1, final_scores=0.5368355,
final_local_scores=0.49832743, pem=-0.6351416,
similarity=22.414797]
13: 1103112 FC_Groningen
```

```
[pem_pos=2, final_scores=0.6567779,
final_local_scores=0.6195752, pem=-1.9677447,
similarity=33.279976]
14: 7632263 Dean_Gorré
    [pem_pos=0, final_scores=0.4925778,
final_local_scores=0.5780459, pem=0.0, similarity=21.339756]
15: 9426084 Ernest_Faber
    [pem_pos=0, final_scores=0.6434843,
final_local_scores=0.6656082, pem=0.0, similarity=23.95279]
16: 1456550 Ronald_Waterreus
    [pem_pos=0, final_scores=0.68385583,
final_local_scores=0.7493216, pem=0.0, similarity=26.450964]
17: 9426084 Ernest_Faber
    [pem_pos=0, final_scores=1.1497655,
final_local_scores=1.2255683, pem=0.0, similarity=40.663116]
18: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.5620763,
final_local_scores=0.50738156, pem=0.0, similarity=19.230991]
19: 1103112 FC_Groningen
    [pem_pos=2, final_scores=0.6264383,
final_local_scores=0.5864338, pem=-1.9677447,
similarity=32.29097]
20: 2945257 Luc_Nilis
    [pem_pos=0, final_scores=0.6261862,
final_local_scores=0.683322, pem=0.0, similarity=24.481405]
21: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.4660508,
final_local_scores=0.4024884, pem=0.0, similarity=16.100769]
22: 2212066 Željko_Petrović
    [pem_pos=0, final_scores=0.6070432,
final_local_scores=0.67876995, pem=0.0, similarity=24.345562]
23: 834256 Vitesse
    [pem_pos=0, final_scores=0.59252286,
final_local_scores=0.55879945, pem=0.0, similarity=20.765404]
24: 834226 FC_Utrecht
    [pem_pos=2, final_scores=0.7334491,
final_local_scores=0.71218246, pem=-2.2398934,
similarity=37.523544]
25: 4287016 Raymond_Atteveld
    [pem_pos=0, final_scores=0.6323675,
final_local_scores=0.70237803, pem=0.0, similarity=25.050074]
26: 5824381 Gaston_Taument
    [pem_pos=0, final_scores=0.5608265,
final_local_scores=0.6243758, pem=0.0, similarity=22.722332]
27: 18951905 Argentina
    [pem_pos=0, final_scores=0.5334715,
final_local_scores=0.651547, pem=-0.18885723,
similarity=24.560211]
28: 9420663 Pablo_Sánchez
    [pem_pos=0, final_scores=0.51936656,
final_local_scores=0.59921026, pem=0.0, similarity=21.971342]
29: 47707 Feyenoord
```

```
[pem_pos=0, final_scores=0.54136807,
final_local_scores=0.49017012, pem=0.0, similarity=18.717367]
30: 2202202 FC_Volendam
    [pem_pos=1, final_scores=0.53263175,
final_local_scores=0.5107504, pem=-0.8958714,
similarity=24.203411]
```

Sample: b'366_SOCCER'

SOCCER - [1 1 BELGIAN 1] FIRST DIVISION RESULTS /
STANDINGS . [2 2 BRUSSELS 2] 1996-08-25 Results of [3 3 Belgian
an 3] first division soccer matches at the weekend : [4 4 Genk
4] 1 [5 5 Club Brugge 5] 1 [6 6 Harelbeke 6] 3 [7 7 Mechelen 7
] 3 [8 8 Standard Liege 8] 3 [9 9 Molenbeek 9] 0 [10 10 Anderlecht 10] 2 [11 11 Lokeren 11] 2 [12 12 Cercle Brugge 12] 2 [1
3 13 Mouscron 13] 2 1 Antwerp 1 [14 14 Lommel 14] 4 [15 15 Ghent 15] 3 [16 16 Aalst 16] 2 [17 17 Lierse 17] 4 [18 18 Charleroi 18] 0 [19 19 Sint Truiden 19] 3 Ekeren 1 Standings
(tabulated under played , won , drawn , lost , goals for ,
against , points) : [20 20 Gent 20] 4 3 1 0 9 5 10 [21 21 Standard Liege 21] 4 3 0 1 7 3 9 [22 22 Club Brugge 22] 4 2 2 0
10 4 8 [23 23 Mouscron 23] 4 2 2 0 7 4 8 [24 24 Anderlecht 24]
4 1 3 0 9 3 6 [25 25 Lierse 25] 4 1 3 0 7 3 6 2 Antwerp 4 2 0
2 6 10 6 [26 26 Genk 26] 4 1 2 1 6 5 5 [27 27 Molenbeek 27] 4
1 2 1 4 5 5 [28 28 Harelbeke 28] 4 1 1 2 6 7 4 [29 29 Aalst 29]
] 4 1 1 2 5 6 4 [30 30 Lokeren 30] 4 1 1 2 4 5 4 Ekeren 4 1 1
2 6 8 4 [31 31 Lommel 31] 4 1 1 2 5 10 4 [32 32 Mechelen 32]
4 0 3 1 6 7 3 [33 33 Cercle Brugge 33] 4 0 3 1 4 5 3 [34 34 Charleroi 34] 4 1 0 3 4 8 3 [35 35 Sint Truiden 35] 4 1 0 3 4
11 3

1: 3343 Belgium
2: 3708 Brussels
3: 3343 Belgium
4: 1433871 K.R.C._Genk
5: 563623 Club_Brugge_K.V.
6: 3462714 K.R.C._Zuid-West-Vlaanderen
7: 564514 KV_Mechelen
8: 990794 Standard_Liège
9: 1452159 R._White_Daring_Molenbeek
10: 26584259 R.S.C._Anderlecht
11: 1437580 K.S.C._Lokeren_Oost-Vlaanderen
12: 1434036 Cercle_Brugge_K.S.V.
13: 1434546 R.E._Mouscron
14: 3470620 K.F.C._Lommel_S.K.
15: 1521107 K.A.A._Gent
16: 1537131 SC_Eendracht_Aalst
17: 842033 Lierse_S.K.
18: 1433845 R._Charleroi_S.C.
19: 1437246 K._Sint-Truidense_V.V.
20: 1521107 K.A.A._Gent
21: 990794 Standard_Liège
22: 563623 Club_Brugge_K.V.
23: 1434546 R.E._Mouscron

24: 26584259 R.S.C._Anderlecht
25: 842033 Lierse_S.K.
26: 1433871 K.R.C._Genk
27: 1452159 R.White_Daring_Molenbeek
28: 3462714 K.R.C._Zuid-West-Vlaanderen
29: 1537131 SC_Eendracht_Aalst
30: 1437580 K.S.C._Lokeren_Oost-Vlaanderen
31: 3470620 K.F.C._Lommel_S.K.
32: 564514 KV_Mechelen
33: 1434036 Cercle_Brugge_K.S.V.
34: 1433845 R.Charleroi_S.C.
35: 1437246 K.Sint-Truidense_V.V.
1: 1520840 Royal_Antwerp_FC
[winner: 2107074 Royal_Antwerp_F.C., pem_pos=2,
final_scores=-0.044693843, final_local_scores=-0.13821405,
pem=-2.9206293, similarity=15.847965]
RECALL MISS: 1520840 Royal_Antwerp_FC]
2: 1520840 Royal_Antwerp_FC
[winner: 2107074 Royal_Antwerp_F.C., pem_pos=2,
final_scores=-0.07831794, final_local_scores=-0.17494327,
pem=-2.9206293, similarity=14.751892]
RECALL MISS: 1520840 Royal_Antwerp_FC]
1: 3343 Belgium
[pem_pos=0, final_scores=0.53929317,
final_local_scores=0.5914323, pem=-0.12403572,
similarity=22.413757]
2: 3708 Brussels
[pem_pos=0, final_scores=0.51587677,
final_local_scores=0.59073305, pem=-0.025440974,
similarity=21.856718]
3: 3343 Belgium
[pem_pos=0, final_scores=0.615513,
final_local_scores=0.6746909, pem=-0.12403572,
similarity=24.89836]
4: 1433871 K.R.C._Genk
[pem_pos=1, final_scores=0.64776784,
final_local_scores=0.61245567, pem=-0.11508816,
similarity=22.99248]
5: 563623 Club_Brugge_K.V.
[pem_pos=0, final_scores=0.4478026,
final_local_scores=0.3874349, pem=0.0, similarity=15.651543]
6: 3462714 K.R.C._Zuid-West-Vlaanderen
[pem_pos=1, final_scores=0.40998638,
final_local_scores=0.43506533, pem=-0.54181004,
similarity=20.019377]
7: 564514 KV_Mechelen
[pem_pos=1, final_scores=0.43004626,
final_local_scores=0.37275726, pem=-1.6237628,
similarity=24.04381]
8: 990794 Standard_Liège
[pem_pos=0, final_scores=0.52603126,
final_local_scores=0.476861, pem=0.0, similarity=18.320196]

```
9: 1452159 R._White_Daring_Molenbeek
    [pem_pos=0, final_scores=0.4231903,
final_local_scores=0.39382046, pem=-0.005970144,
similarity=15.874569]
10: 26584259 R.S.C._Anderlecht
    [pem_pos=0, final_scores=0.637147,
final_local_scores=0.59415376, pem=0.0, similarity=21.820446]
11: 1437580 K.S.C._Lokeren_Oost-Vlaanderen
    [pem_pos=0, final_scores=0.5553923,
final_local_scores=0.51682305, pem=0.0, similarity=19.512743]
12: 1434036 Cercle_Brugge_K.S.V.
    [pem_pos=0, final_scores=0.5082827,
final_local_scores=0.45909864, pem=0.0, similarity=17.790133]
13: 1434546 R.E._Mouscron
    [pem_pos=1, final_scores=0.49385554,
final_local_scores=0.47845906, pem=-0.28243667,
similarity=19.90382]
14: 3470620 K.F.C._Lommel_S.K.
    [pem_pos=1, final_scores=0.438801,
final_local_scores=0.4509439, pem=-1.6453731,
similarity=26.494572]
15: 1521107 K.A.A._Gent
    [pem_pos=13, final_scores=0.5511747,
final_local_scores=0.5109264, pem=-2.9399738,
similarity=35.32481]
16: 1537131 SC_Eendracht_Aalst
    [pem_pos=1, final_scores=0.5028804,
final_local_scores=0.47750682, pem=-1.4374129,
similarity=26.156343]
17: 842033 Lierse_S.K.
    [pem_pos=0, final_scores=0.5168749,
final_local_scores=0.47142482, pem=0.0, similarity=18.15797]
18: 1433845 R.Charleroi_S.C.
    [pem_pos=1, final_scores=0.5918631,
final_local_scores=0.5605022, pem=-1.0632312,
similarity=26.598232]
19: 1437246 K._Sint-Truidense_V.V.
    [pem_pos=0, final_scores=0.570849,
final_local_scores=0.5427466, pem=0.0, similarity=20.286354]
20: 1521107 K.A.A._Gent
    [pem_pos=13, final_scores=0.5676332,
final_local_scores=0.52890486, pem=-2.9399738,
similarity=35.861324]
21: 990794 Standard_Liège
    [pem_pos=0, final_scores=0.51084644,
final_local_scores=0.46027392, pem=0.0, similarity=17.825205]
22: 563623 Club_Brugge_K.V.
    [pem_pos=0, final_scores=0.4538749,
final_local_scores=0.394068, pem=0.0, similarity=15.849488]
23: 1434546 R.E._Mouscron
    [pem_pos=1, final_scores=0.4890644,
final_local_scores=0.47322547, pem=-0.28243667,
```

```
similarity=19.74764]
24: 26584259 R.S.C._Anderlecht
    [pem_pos=0, final_scores=0.6135048,
final_local_scores=0.56832826, pem=0.0, similarity=21.049763]
25: 842033 Lierse_S.K.
    [pem_pos=0, final_scores=0.5254026,
final_local_scores=0.48074007, pem=0.0, similarity=18.435955]
26: 1433871 K.R.C._Genk
    [pem_pos=1, final_scores=0.66320235,
final_local_scores=0.62931544, pem=-0.11508816,
similarity=23.49561]
27: 1452159 R._White_Daring_Molenbeek
    [pem_pos=0, final_scores=0.4503023,
final_local_scores=0.42343622, pem=-0.005970144,
similarity=16.758362]
28: 3462714 K.R.C._Zuid-West-Vlaanderen
    [pem_pos=1, final_scores=0.47314638,
final_local_scores=0.50405806, pem=-0.54181004,
similarity=22.078259]
29: 1537131 SC_Eendracht_Aalst
    [pem_pos=1, final_scores=0.47106987,
final_local_scores=0.44275868, pem=-1.4374129,
similarity=25.119389]
30: 1437580 K.S.C._Lokeren_Oost-Vlaanderen
    [pem_pos=0, final_scores=0.619902,
final_local_scores=0.5872901, pem=0.0, similarity=21.615622]
31: 3470620 K.F.C._Lommel_S.K.
    [pem_pos=1, final_scores=0.39001757,
final_local_scores=0.39765543, pem=-1.6453731,
similarity=24.904339]
32: 564514 KV_Mechelen
    [pem_pos=1, final_scores=0.41868407,
final_local_scores=0.36034575, pem=-1.6237628,
similarity=23.673424]
33: 1434036 Cercle_Brugge_K.S.V.
    [pem_pos=0, final_scores=0.5194734,
final_local_scores=0.47132272, pem=0.0, similarity=18.154922]
34: 1433845 R._Charleroi_S.C.
    [pem_pos=1, final_scores=0.4567485,
final_local_scores=0.41290998, pem=-1.0632312,
similarity=22.193787]
35: 1437246 K._Sint-Truidense_V.V.
    [pem_pos=0, final_scores=0.5234657,
final_local_scores=0.49098748, pem=0.0, similarity=18.741758]
```

Sample: b'367_SOCCER'
SOCCER - [1 1 DUTCH 1] FIRST DIVISION SUMMARIES . [2 2 AMSTE
RDAM 2] 1996-08-25 Summary of [1 1 Dutch 1] first division
soccer played on Sunday : [3 3 Feyenoord Rotterdam 3] 3
(Sanchez 27th , [4 4 Taument 4] 44th , 57th) [5 5 Volendam 5
6] 0 [7 7 AZ Alkmaar 7] 0 . Attendance not given . [8 8 Vites

se Arnhem 8] 1 (Van Wanrooy 58th) [9 9 Utrecht 9] 0 . Halftime 0-0 . Attendance 7,032 . [10 10 Twente Enschede 10] 1 ([11 11 Hoogma 11] 30th) [12 12 Roda JC Kerkrade 12] 1 ([13 13 Roelofsen 13] 28th) . Halftime 1-1 . Attendance not given . [14 14 PSV Eindhoven 14] 4 ([15 15 Faber 15] 32nd , [16 16 Nilis 16] 63rd 79th , Petrovic 78th) [17 17 Groningen 1 7] 1 (Gorre 7th) . Halftime 1-1 . Attendance 27,500 Played on Saturday : Graafschap Doetinchem 3 (Ibrahim 20th 63rd , [1 8 18 Godee 18] 54th) [19 19 RKC Waalwijk 19] 2 (Dos Santos 38th , Van Arum 73th penalty) . Halftime 1-1 . Attendance 7,000 [20 20 Willem II Tilburg 20] 0 [21 21 Fortuna Sittard 21] 1 ([22 22 Hamming 22] 65th) . Halftime 0-0 . Attendance 7,250 . [23 23 NAC Breda 23] 1 (Arnold 70th) [24 24 Sparta Rotterdam 24] 0 . Halftime 0-0 . Attendance 11,500 . [25 25 Heerenveen 25] 2 (Wouden 53rd , [26 26 Dahl Tomasson 26] 74th) [27 27 Ajax Amsterdam 27] 0 . Halftime 0-0 . Attendance 13,500 .

1: 21148 Netherlands
2: 844 Amsterdam
3: 47707 Feyenoord
4: 5824381 Gaston_Taument
5: 2202202 FC_Volendam
6: 834154 N.E.C._(football.club)
7: 2289 AZ_Alkmaar
8: 834256 Vitesse
9: 834226 FC_Utrecht
10: 800033 FC_Twente
11: 17554690 Nico-Jan_Hoogma
12: 834196 Roda_JC_Kerkrade
13: 9423088 Richard_Roelofsen
14: 79820 PSV_Eindhoven
15: 9426084 Ernest_Faber
16: 2945257 Luc_Nilis
17: 1103112 FC_Groningen
18: 20868248 Edwin_Godee
19: 1107824 RKC_Waalwijk
20: 834275 Willem_II_(football.club)
21: 1336325 Fortuna_Sittard
22: 9306596 Ronald_Hamming
23: 505173 NAC_Breda
24: 79818 Sparta_Rotterdam
25: 745757 SC_Heerenveen
26: 735169 Jon_Dahl_Tomasson
27: 2273 AFC_Ajax

1: 1332105 Eredivisie
[winner: 21148 Netherlands, pem_pos=0,
final_scores=0.39580083, final_local_scores=0.42929322,
pem=-0.7943452, similarity=21.220451]
RECALL MISS: 1332105 Eredivisie]

1: 21148 Netherlands
[pem_pos=0, final_scores=0.6315227,
final_local_scores=0.68678355, pem=-0.7943452,

```
similarity=28.904476]
2: 844 Amsterdam
    [pem_pos=0, final_scores=0.5773953,
final_local_scores=0.63753027, pem=-0.0642803,
similarity=23.464455]
3: 47707 Feyenoord
    [pem_pos=0, final_scores=0.63810724,
final_local_scores=0.59536934, pem=0.0, similarity=21.856722]
4: 5824381 Gaston_Taument
    [pem_pos=0, final_scores=0.56602395,
final_local_scores=0.62940335, pem=0.0, similarity=22.872364]
5: 2202202 FC_Volendam
    [pem_pos=1, final_scores=0.68332577,
final_local_scores=0.6670926, pem=-0.8958714,
similarity=28.868975]
6: 834154 N.E.C._(football_club)
    [pem_pos=0, final_scores=0.6349372,
final_local_scores=0.59855, pem=0.0, similarity=21.95164]
7: 2289 AZ_Alkmaar
    [pem_pos=0, final_scores=0.49807853,
final_local_scores=0.44469863, pem=0.0, similarity=17.360407]
8: 834256 Vitesse
    [pem_pos=0, final_scores=0.56207675,
final_local_scores=0.5220434, pem=0.0, similarity=19.66853]
9: 834226 FC_Utrecht
    [pem_pos=2, final_scores=0.79644483,
final_local_scores=0.7763645, pem=-2.2398934,
similarity=39.438866]
10: 800033 FC_Twente
    [pem_pos=0, final_scores=0.56206745,
final_local_scores=0.5176565, pem=0.0, similarity=19.537615]
11: 17554690 Nico-Jan_Hoogma
    [pem_pos=0, final_scores=0.578433,
final_local_scores=0.6567021, pem=0.0, similarity=23.687014]
12: 834196 Roda_JC_Kerkrade
    [pem_pos=0, final_scores=0.6315228,
final_local_scores=0.6045589, pem=0.0, similarity=22.130957]
13: 9423088 Richard_Roelofsen
    [pem_pos=1, final_scores=0.30637056,
final_local_scores=0.35981756, pem=-0.6931472,
similarity=18.596828]
14: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.53487355,
final_local_scores=0.4841249, pem=0.0, similarity=18.536966]
15: 9426084 Ernest_Faber
    [pem_pos=9, final_scores=0.5996865,
final_local_scores=0.6378518, pem=-3.4590578,
similarity=41.935375]
16: 2945257 Luc_Nilis
    [pem_pos=0, final_scores=0.65388626,
final_local_scores=0.72531575, pem=0.0, similarity=25.734581]
17: 1103112 FC_Groningen
```

```

[pem_pos=2, final_scores=0.81149787,
final_local_scores=0.78635186, pem=-1.9677447,
similarity=38.256924]
18: 20868248 Edwin_Godee
    [pem_pos=1, final_scores=0.29312557,
final_local_scores=0.35681242, pem=-0.6931472,
similarity=18.507149]
19: 1107824 RKC_Waalwijk
    [pem_pos=0, final_scores=0.59122276,
final_local_scores=0.55792147, pem=0.0, similarity=20.739202]
20: 834275 Willem_II_(football_club)
    [pem_pos=0, final_scores=0.7503142,
final_local_scores=0.7231383, pem=0.0, similarity=25.669601]
21: 1336325 Fortuna_Sittard
    [pem_pos=0, final_scores=0.8756749,
final_local_scores=0.8728866, pem=0.0, similarity=30.13839]
22: 9306596 Ronald_Hamming
    [pem_pos=1, final_scores=0.56151587,
final_local_scores=0.64856714, pem=-0.6931472,
similarity=27.213692]
23: 505173 NAC_Breda
    [pem_pos=0, final_scores=0.6115303,
final_local_scores=0.57901967, pem=0.0, similarity=21.368814]
24: 79818 Sparta_Rotterdam
    [pem_pos=0, final_scores=0.49302393,
final_local_scores=0.4513905, pem=0.0, similarity=17.560106]
25: 745757 SC_Heerenveen
    [pem_pos=1, final_scores=0.5775487,
final_local_scores=0.53634125, pem=-0.6351416,
similarity=23.549204]
26: 735169 Jon_Dahl_Tomasson
    [pem_pos=0, final_scores=0.737022,
final_local_scores=0.7895758, pem=0.0, similarity=27.65223]
27: 2273 AFC_Ajax
    [pem_pos=0, final_scores=0.627238,
final_local_scores=0.5890771, pem=0.0, similarity=21.66895]
1: 21148 Netherlands
    [pem_pos=0, final_scores=0.39580083,
final_local_scores=0.42929322, pem=-0.7943452,
similarity=21.220451]

```

Sample: b'368_SOCCER'

```

SOCCER - [1 1 DUTCH 1] FIRST DIVISION RESULTS / TABLE . [2 2
AMSTERDAM 2] 1996-08-25 Result of [3 3 Dutch 3] first
division soccer match played on Sunday : [4 4 Feyenoord Rotte
rdam 4] 3 [5 5 Volendam 5] 0 [6 6 NEC Nijmegen 6] 0 [7 7 AZ Al
kmaar 7] 0 [8 8 Vitesse Arnhem 8] 1 [9 9 Utrecht 9] 0 [10 10 T
wente Enschede 10] 1 [11 11 Roda JC 11] 1 [12 12 PSV Eindhoven
12] 4 [13 13 Groningen 13] 1 Played on Saturday : Graafschap
Doetinchem 3 [14 14 RKC Waalwijk 14] 2 [15 15 Willem II Tilbu
rg 15] 0 [16 16 Fortuna Sittard 16] 1 [17 17 NAC Breda 17] 1 [
18 18 Sparta Rotterdam 18] 0 [19 19 Heerenveen 19] 2 [20 20 Aj

```

ax Amsterdam 20] 0 Standings (tabulate under played , won ,
drawn , lost , goals for , goals against , points) : [21 21 P
SV Eindhoven 21] 2 2 0 0 8 2 6 [22 22 Vitesse Arnhem 22] 2 2
0 0 3 0 6 [23 23 Feyenoord Rotterdam 23] 2 1 1 0 4 1 4
Graafschap Doetinchem 2 1 1 0 4 3 4 [24 24 Twente Enschede 24]
2 1 1 0 4 2 4 [25 25 Fortuna Sittard 25] 2 1 1 0 1 0 4 [26 26
Heerenveen 26] 2 1 0 1 3 3 3 [27 27 NAC Breda 27] 2 1 0 1 1 1
3 [28 28 Ajax Amsterdam 28] 2 1 0 1 1 2 3 [29 29 Roda JC Kerkrade 29] 2 0 2 0 2 2 2 [30 30 Utrecht 30] 2 0 1 1 2 3 1 [31 31
RKC Waalwijk 31] 2 0 1 1 4 5 1 [32 32 Sparta Rotterdam 32] 2
0 1 1 0 1 1 [33 33 Willem II Tilburg 33] 2 0 1 1 0 1 1 [34 34
AZ Alkmaar 34] 2 0 1 1 0 2 1 [35 35 Volendam 35] 2 0 1 1 1 4
1 [36 36 Groningen 36] 2 0 1 1 1 4 1 [37 37 NEC Nijmegen 37]
2 0 1 1 1 4 1
1: 21148 Netherlands
2: 844 Amsterdam
3: 21148 Netherlands
4: 47707 Feyenoord
5: 2202202 FC_Volendam
6: 834154 N.E.C._(football_club)
7: 2289 AZ_Alkmaar
8: 834256 Vitesse
9: 834226 FC_Utrecht
10: 800033 FC_Twente
11: 834196 Roda_JC_Kerkrade
12: 79820 PSV_Eindhoven
13: 1103112 FC_Groningen
14: 1107824 RKC_Waalwijk
15: 834275 Willem_II_(football_club)
16: 1336325 Fortuna_Sittard
17: 505173 NAC_Breda
18: 79818 Sparta_Rotterdam
19: 745757 SC_Herenveen
20: 2273 AFC_Ajax
21: 79820 PSV_Eindhoven
22: 834256 Vitesse
23: 47707 Feyenoord
24: 800033 FC_Twente
25: 1336325 Fortuna_Sittard
26: 745757 SC_Herenveen
27: 505173 NAC_Breda
28: 2273 AFC_Ajax
29: 834196 Roda_JC_Kerkrade
30: 834226 FC_Utrecht
31: 1107824 RKC_Waalwijk
32: 79818 Sparta_Rotterdam
33: 834275 Willem_II_(football_club)
34: 2289 AZ_Alkmaar
35: 2202202 FC_Volendam
36: 1103112 FC_Groningen
37: 834154 N.E.C._(football_club)
1: 21148 Netherlands

```
[pem_pos=0, final_scores=0.6410521,
final_local_scores=0.70662075, pem=-0.7943452,
similarity=29.496458]
2: 844 Amsterdam
    [pem_pos=0, final_scores=0.57126427,
final_local_scores=0.6379909, pem=-0.0642803,
similarity=23.4782]
3: 21148 Netherlands
    [pem_pos=0, final_scores=0.46397644,
final_local_scores=0.5131925, pem=-0.7943452,
similarity=23.724173]
4: 47707 Feyenoord
    [pem_pos=0, final_scores=0.6472926,
final_local_scores=0.6004541, pem=0.0, similarity=22.008461]
5: 2202202 FC_Volendam
    [pem_pos=1, final_scores=0.6525745,
final_local_scores=0.62730926, pem=-0.8958714,
similarity=27.68176]
6: 834154 N.E.C._(football_club)
    [pem_pos=0, final_scores=0.6332055,
final_local_scores=0.59054595, pem=0.0, similarity=21.712782]
7: 2289 AZ_Alkmaar
    [pem_pos=0, final_scores=0.4971128,
final_local_scores=0.43713802, pem=0.0, similarity=17.134783]
8: 834256 Vitesse
    [pem_pos=0, final_scores=0.5775575,
final_local_scores=0.5323736, pem=0.0, similarity=19.976803]
9: 834226 FC_Utrecht
    [pem_pos=2, final_scores=0.7198761,
final_local_scores=0.68763083, pem=-2.2398934,
similarity=36.790874]
10: 800033 FC_Twente
    [pem_pos=0, final_scores=0.5448683,
final_local_scores=0.49154365, pem=0.0, similarity=18.758356]
11: 834196 Roda_JC_Kerkrade
    [pem_pos=0, final_scores=0.50374776,
final_local_scores=0.45947373, pem=0.0, similarity=17.801325]
12: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.48777527,
final_local_scores=0.42670196, pem=0.0, similarity=16.82335]
13: 1103112 FC_Groningen
    [pem_pos=2, final_scores=0.6970225,
final_local_scores=0.65648323, pem=-1.9677447,
similarity=34.381386]
14: 1107824 RKC_Waalwijk
    [pem_pos=0, final_scores=0.5993916,
final_local_scores=0.56089216, pem=0.0, similarity=20.827854]
15: 834275 Willem_II_(football_club)
    [pem_pos=0, final_scores=0.78820354,
final_local_scores=0.7579392, pem=0.0, similarity=26.708132]
16: 1336325 Fortuna_Sittard
    [pem_pos=0, final_scores=0.7320864,
```

```
final_local_scores=0.70821685, pem=0.0, similarity=25.224316]
17: 505173 NAC_Breda
    [pem_pos=0, final_scores=0.6692092,
final_local_scores=0.63405675, pem=0.0, similarity=23.01123]
18: 79818 Sparta_Rotterdam
    [pem_pos=0, final_scores=0.57156813,
final_local_scores=0.53183854, pem=0.0, similarity=19.960836]
19: 745757 SC_Heerenveen
    [pem_pos=1, final_scores=0.6496442,
final_local_scores=0.6086901, pem=-0.6351416,
similarity=25.708239]
20: 2273 AFC_Ajax
    [pem_pos=0, final_scores=0.67783517,
final_local_scores=0.6397207, pem=0.0, similarity=23.180254]
21: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.5030571,
final_local_scores=0.44339508, pem=0.0, similarity=17.321507]
22: 834256 Vitesse
    [pem_pos=0, final_scores=0.5565409,
final_local_scores=0.50941616, pem=0.0, similarity=19.291708]
23: 47707 Feyenoord
    [pem_pos=0, final_scores=0.5706189,
final_local_scores=0.5166997, pem=0.0, similarity=19.509062]
24: 800033 FC_Twente
    [pem_pos=0, final_scores=0.5268426,
final_local_scores=0.47185332, pem=0.0, similarity=18.170757]
25: 1336325 Fortuna_Sittard
    [pem_pos=0, final_scores=0.72431237,
final_local_scores=0.69972485, pem=0.0, similarity=24.970898]
26: 745757 SC_Heerenveen
    [pem_pos=1, final_scores=0.6304863,
final_local_scores=0.587763, pem=-0.6351416,
similarity=25.08373]
27: 505173 NAC_Breda
    [pem_pos=0, final_scores=0.6763943,
final_local_scores=0.6419054, pem=0.0, similarity=23.245451]
28: 2273 AFC_Ajax
    [pem_pos=0, final_scores=0.61776626,
final_local_scores=0.57410455, pem=0.0, similarity=21.222137]
29: 834196 Roda_JC_Kerkrade
    [pem_pos=0, final_scores=0.60851574,
final_local_scores=0.5739168, pem=0.0, similarity=21.216536]
30: 834226 FC_Utrecht
    [pem_pos=2, final_scores=0.7043014,
final_local_scores=0.67061776, pem=-2.2398934,
similarity=36.283173]
31: 1107824 RKC_Waalwijk
    [pem_pos=0, final_scores=0.5550265,
final_local_scores=0.51243, pem=0.0, similarity=19.381647]
32: 79818 Sparta_Rotterdam
    [pem_pos=0, final_scores=0.6009874,
final_local_scores=0.5639746, pem=0.0, similarity=20.91984]
```

```
33: 834275 Willem_II_(football_club)
    [pem_pos=0, final_scores=0.7395087,
final_local_scores=0.70474744, pem=0.0, similarity=25.120783]
34: 2289 AZ_Alkmaar
    [pem_pos=0, final_scores=0.48114496,
final_local_scores=0.41969556, pem=0.0, similarity=16.614265]
35: 2202202 FC_Volendam
    [pem_pos=1, final_scores=0.69313353,
final_local_scores=0.67161393, pem=-0.8958714,
similarity=29.0039]
36: 1103112 FC_Groningen
    [pem_pos=2, final_scores=0.7452387,
final_local_scores=0.70915216, pem=-1.9677447,
similarity=35.95313]
37: 834154 N.E.C._(football_club)
    [pem_pos=0, final_scores=0.6217999,
final_local_scores=0.57808703, pem=0.0, similarity=21.340982]
```

Sample: b'369_DUTCH'

```
[1 1 DUTCH 1] CAPTAIN [2 2 BLIND 2] ENDS INTERNATIONAL CAREER . [3 3 AMSTERDAM 3] 1996-08-25 [4 4 Dutch 4] soccer captain [5 5 Danny Blind 5] has decided to end his international career , [6 6 Ajax 6] spokesman David Endt said on Sunday . Endt told [7 7 Dutch 7] news agency [8 8 ANP 8] that [9 9 Blind 9] , 35 , would no longer be available for selection for the national squad . The [10 10 Ajax 10] defender , who led the [11 11 Netherlands 11] into the quarter-finals at June 's [12 12 European 12] championship finals in [13 13 England 13] , had decided to devote his attention to playing for his [14 14 Amsterdam 14] club , Endt said . [15 15 Blind 15] , who played in the [16 16 1990 World Cup 16] and the [17 17 1992 European championship 17] , was capped 42 times for the [18 18 Netherlands 18] .
```

```
1: 21148 Netherlands
2: 821908 Danny_Blind
3: 844 Amsterdam
4: 21148 Netherlands
5: 821908 Danny_Blind
6: 2273 AFC_Ajax
7: 21148 Netherlands
8: 2161974 Algemeen_Nederlands_Persbureau
9: 821908 Danny_Blind
10: 2273 AFC_Ajax
11: 9647657 Netherlands_national_football_team
12: 9239 Europe
13: 9316 England
14: 844 Amsterdam
15: 821908 Danny_Blind
16: 157232 1990_FIFA_World_Cup
17: 558990 UEFA_Euro_1992
18: 9647657 Netherlands_national_football_team
1: 21148 Netherlands
```

```
[pem_pos=0, final_scores=0.3820029,
final_local_scores=0.3617261, pem=-0.7943452,
similarity=19.204115]
2: 821908 Danny_Blind
    [pem_pos=0, final_scores=0.76217896,
final_local_scores=0.77554524, pem=0.0, similarity=27.23353]
3: 844 Amsterdam
    [pem_pos=0, final_scores=0.58688897,
final_local_scores=0.59530294, pem=-0.0642803,
similarity=22.204308]
4: 21148 Netherlands
    [pem_pos=0, final_scores=0.49564117,
final_local_scores=0.48585862, pem=-0.7943452,
similarity=22.908476]
5: 821908 Danny_Blind
    [pem_pos=0, final_scores=0.9156,
final_local_scores=0.9352513, pem=0.0, similarity=31.999475]
6: 2273 AFC_Ajax
    [pem_pos=1, final_scores=0.5417865,
final_local_scores=0.52056855, pem=-1.4356393,
similarity=27.431747]
7: 21148 Netherlands
    [pem_pos=0, final_scores=0.58889204,
final_local_scores=0.5877211, pem=-0.7943452,
similarity=25.948256]
8: 2161974 Algemeen_Nederlands_Persbureau
    [pem_pos=1, final_scores=0.42588574,
final_local_scores=0.59741455, pem=-0.9606266,
similarity=27.141792]
9: 821908 Danny_Blind
    [pem_pos=0, final_scores=1.0222819,
final_local_scores=1.0596683, pem=0.0, similarity=35.71233]
10: 2273 AFC_Ajax
    [pem_pos=1, final_scores=0.57935876,
final_local_scores=0.5616105, pem=-1.4356393,
similarity=28.65652]
11: 9647657 Netherlands_national_football_team
    [pem_pos=1, final_scores=0.5246213,
final_local_scores=0.5093503, pem=-2.660505,
similarity=33.75798]
12: 9239 Europe
    [pem_pos=0, final_scores=0.5346536,
final_local_scores=0.6197917, pem=-0.25557294,
similarity=23.975378]
13: 9316 England
    [pem_pos=0, final_scores=0.4936288,
final_local_scores=0.5650629, pem=-0.19053856,
similarity=21.988495]
14: 844 Amsterdam
    [pem_pos=0, final_scores=0.61908776,
final_local_scores=0.6304753, pem=-0.0642803,
similarity=23.25392]
```

```
15: 821908 Danny_Blind
    [pem_pos=0, final_scores=1.035567,
final_local_scores=1.0741804, pem=0.0, similarity=36.1454]
16: 157232 1990_FIFA_World_Cup
    [pem_pos=0, final_scores=0.1658781,
final_local_scores=0.14858946, pem=0.0, similarity=8.523919]
17: 558990 UEFA_Euro_1992
    [pem_pos=0, final_scores=0.35119087,
final_local_scores=0.356779, pem=0.0, similarity=14.736711]
18: 9647657 Netherlands_national_football_team
    [pem_pos=1, final_scores=0.44011325,
final_local_scores=0.41703808, pem=-2.660505,
similarity=31.003202]
```

Sample: b'370_CRICKET'

CRICKET - [1 1 INDIA 1] BANS [2 2 SIDHU 2] FOR 50 DAYS . [3 3 NEW DELHI 3] 1996-08-25 [4 4 Indian 4] opener [5 5 Navjot Singh Sidhu 5] was on Sunday given a 50-day ban from international cricket for quitting this year 's tour of [6 6 England 6] , the [7 7 Press Trust of India 7] said . The right-handed batsman will have to forfeit half the money he was due to earn from the tour , the news agency said after a disciplinary committee set up by the [8 8 Board of Control for Cricket in India 8] met at [9 9 Mohali 9] , near the northern city of [10 10 Chandigarh 10] . [11 11 Sidhu 11] abandoned the [12 12 Indian 12] team after the third one-day international against [13 13 England 13] at [14 14 Old Trafford 14] in [15 15 Manchester 15] on May 26 , before [16 16 India 16] began a three-test series , citing serious differences with captain [17 17 Mohammed Azharuddin 17] . [18 18 Azharuddin 18] was sacked after the tour and replaced by [19 19 Sachin Tendulkar 19] . [20 20 Sidhu 20] was not considered for the four-nation Singer Cup beginning in [21 21 Sri Lanka 21] this month and the Sahara Cup against [22 22 Pakistan 22] scheduled to be played in [23 23 Canada 23] next month . [24 24 Sidhu 24] , whose ban ends on October 14 , will be free to play domestic cricket . He will not be considered for a test match against [25 25 Australia 25] starting on October 10 in [26 26 New Delhi 26] , the [27 27 United News of India 27] said .

```
1: 407754 India_national_cricket_team
2: 583231 Navjot_Singh_Sidhu
3: 51585 New_Delhi
4: 14533 India
5: 583231 Navjot_Singh_Sidhu
6: 9316 England
7: 1724500 Press_Trust_of_India
8: 564006 Board_of_Control_for_Cricket_in_India
9: 1329327 Ajitgarh
10: 211915 Chandigarh
11: 583231 Navjot_Singh_Sidhu
12: 14533 India
```

```
13: 759125 England_cricket_team
14: 660597 Old_Trafford_Cricket_Ground
15: 20206 Manchester
16: 407754 India_national_cricket_team
17: 86690 Mohammad_Azharuddin
18: 86690 Mohammad_Azharuddin
19: 57570 Sachin_Tendulkar
20: 583231 Navjot_Singh_Sidhu
21: 26750 Sri_Lanka
22: 759295 Pakistan_national_cricket_team
23: 5042916 Canada
24: 583231 Navjot_Singh_Sidhu
25: 374833 Australia_national_cricket_team
26: 51585 New_Delhi
27: 1724515 United_News_of_India
1: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.65797204,
final_local_scores=0.64855874, pem=-3.5152702,
similarity=42.560585]
2: 583231 Navjot_Singh_Sidhu
    [pem_pos=0, final_scores=1.025579,
final_local_scores=1.0560763, pem=0.0, similarity=35.605133]
3: 51585 New_Delhi
    [pem_pos=0, final_scores=0.4019968,
final_local_scores=0.41643083, pem=0.0, similarity=16.51684]
4: 14533 India
    [pem_pos=0, final_scores=0.55665636,
final_local_scores=0.5852866, pem=-0.618294,
similarity=24.918213]
5: 583231 Navjot_Singh_Sidhu
    [pem_pos=0, final_scores=0.50019485,
final_local_scores=0.47722667, pem=0.0, similarity=18.331108]
6: 9316 England
    [pem_pos=0, final_scores=0.4588037,
final_local_scores=0.5091102, pem=-0.19053856,
similarity=20.318756]
7: 1724500 Press_Trust_of_India
    [pem_pos=0, final_scores=0.5693889,
final_local_scores=0.6323842, pem=0.0, similarity=22.961319]
8: 564006 Board_of_Control_for_Cricket_in_India
    [pem_pos=0, final_scores=0.544753,
final_local_scores=0.560113, pem=0.0, similarity=20.804602]
9: 1329327 Ajitgarh
    [pem_pos=0, final_scores=0.26997077,
final_local_scores=0.30741993, pem=0.0, similarity=13.263737]
10: 211915 Chandigarh
    [pem_pos=0, final_scores=0.38461065,
final_local_scores=0.40098912, pem=0.0, similarity=16.056028]
11: 583231 Navjot_Singh_Sidhu
    [pem_pos=0, final_scores=1.0125102,
final_local_scores=1.0418007, pem=0.0, similarity=35.179123]
12: 14533 India
```

```
[pem_pos=0, final_scores=0.66413605,
final_local_scores=0.70269185, pem=-0.618294,
similarity=28.42182]
13: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.6150052,
final_local_scores=0.624281, pem=-3.3054674,
similarity=40.69515]
14: 660597 Old_Trafford_Cricket_Ground
    [pem_pos=1, final_scores=0.54436773,
final_local_scores=0.5899871, pem=-0.6437659,
similarity=25.197004]
15: 20206 Manchester
    [pem_pos=0, final_scores=0.51422775,
final_local_scores=0.574683, pem=-0.3209542,
similarity=22.984798]
16: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.675582,
final_local_scores=0.66779494, pem=-3.5152702,
similarity=43.134632]
17: 86690 Mohammad_Azharuddin
    [pem_pos=0, final_scores=0.60490054,
final_local_scores=0.62428075, pem=0.0, similarity=22.719496]
18: 86690 Mohammad_Azharuddin
    [pem_pos=0, final_scores=0.3541668,
final_local_scores=0.3503923, pem=0.0, similarity=14.546119]
19: 57570 Sachin_Tendulkar
    [pem_pos=0, final_scores=0.54844284,
final_local_scores=0.5447271, pem=0.0, similarity=20.345455]
20: 583231 Navjot_Singh_Sidhu
    [pem_pos=0, final_scores=1.107314,
final_local_scores=1.1453595, pem=0.0, similarity=38.269524]
21: 26750 Sri_Lanka
    [pem_pos=0, final_scores=0.42979807,
final_local_scores=0.46060175, pem=-0.3027737,
similarity=19.481518]
22: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7372458,
final_local_scores=0.7503417, pem=-2.7556365,
similarity=41.466984]
23: 5042916 Canada
    [pem_pos=0, final_scores=0.4718842,
final_local_scores=0.54229325, pem=-0.11662793,
similarity=20.907066]
24: 583231 Navjot_Singh_Sidhu
    [pem_pos=0, final_scores=1.1330304,
final_local_scores=1.1734507, pem=0.0, similarity=39.107822]
25: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.72354037,
final_local_scores=0.74088824, pem=-3.4989398,
similarity=45.227074]
26: 51585 New_Delhi
    [pem_pos=0, final_scores=0.40408498,
```

```
final_local_scores=0.41871184, pem=0.0, similarity=16.58491]
27: 1724515 United_News_of_India
    [pem_pos=0, final_scores=0.4917844,
final_local_scores=0.6284813, pem=0.0, similarity=22.844849]
```

Sample: b'371_RUGBY'

```
[1 1 RUGBY LEAGUE 1] - [1 1 Australian 1] rugby league
standings . [2 2 SYDNEY 2] 1996-08-26 [3 3 Australian 3]
rugby league premiership standings after matches played at
the weekend ( tabulate under played , won , drawn , lost ,
points for , against , total points ) : [4 4 Manly 4] 21 17 0
4 501 181 34 [5 5 Brisbane 5] 21 16 0 5 569 257 32 [6 6 North
Sydney 6] 21 14 2 5 560 317 30 [7 7 Sydney City 7] 20 14 1 5
487 293 29 [8 8 Cronulla 8] 20 12 2 6 359 258 26 [9 9 Canberra
9] 21 12 1 8 502 374 25 [2 2 St George 2] 21 12 1 8 421 344
25 [10 10 Newcastle 10] 21 11 1 9 416 366 23 [11 11 Western Su
burbs 11] 21 11 1 9 382 426 23 [12 12 Auckland 12] 21 11 0 10
406 389 22 [13 13 Sydney Tigers 13] 21 11 0 10 309 435 22 [14
14 Parramatta 14] 21 10 1 10 388 391 21 [15 15 Sydney Bulldogs
15] 21 10 0 11 325 356 20 [16 16 Illawarra 16] 21 8 0 13 395
432 16 [17 17 Western Reds 17] 21 6 1 14 297 398 13 [18 18 Pe
nrith 18] 21 6 1 14 339 448 13 [19 19 North Queensland 19] 21
6 0 15 266 593 12 [20 20 Gold Coast 20] 21 5 1 15 351 483 11 [
21 21 South Sydney 21] 21 5 1 15 304 586 11 [22 22 South Quee
nsland 22] 21 4 0 17 210 460 8 -- Sydney Newsroom 61-2
9373-1800
```

```
1: 4689264 Australia
2: 27862 Sydney
3: 4689264 Australia
4: 885301 Manly-Warringah_Sea_Eagles
5: 5025 Brisbane_Broncos
6: 693234 North_Sydney_Bears
7: 424878 Sydney_Roosters
8: 525034 Cronulla-Sutherland_Sharks
9: 454452 Canberra_Raiders
10: 797018 Newcastle_Knights
11: 1532351 Western_Suburbs_Magpies
12: 702650 New_Zealand_Warriors
13: 1464504 Balmain_Tigers
14: 869062 Parramatta_Eels
15: 516739 Canterbury-Bankstown_Bulldogs
16: 853881 Illawarra_Steelers
17: 2015377 WA_Reds
18: 471314 Penrith_Panthers
19: 995013 North_Queensland_Cowboys
20: 1613020 Gold_Coast_Titans
21: 1079930 South_Sydney_Rabbitohs
22: 2444787 South_Queensland_Crushers
1: 25402 Rugby_football
    [winner: 25735 Rugby_league, pem_pos=0,
final_scores=0.08399956, final_local_scores=0.033355594,
pem=0.0, similarity=5.08511|
```

RECALL MISS: 25402 Rugby_football]
2: 4286852 St. George_Illawarra_Dragons
[winner: 853089 St._George_Dragons, pem_pos=2,
final_scores=0.3669625, final_local_scores=0.30818453,
pem=-2.12389, similarity=24.836597]
gt: 4286852 St._George_Illawarra_Dragons, pem_pos=18,
final_scores=0.16707845, final_local_scores=0.087153435,
pem=-3.28727, similarity=24.567232]
1: 4689264 Australia
[pem_pos=0, final_scores=0.535935,
final_local_scores=0.6130032, pem=-0.491009,
similarity=25.053133]
2: 27862 Sydney
[pem_pos=0, final_scores=0.50873107,
final_local_scores=0.5651787, pem=-0.1443489,
similarity=21.740765]
3: 4689264 Australia
[pem_pos=0, final_scores=0.5392083,
final_local_scores=0.61657876, pem=-0.491009,
similarity=25.159836]
4: 885301 Manly-Warringah_Sea_Eagles
[pem_pos=1, final_scores=0.51518863,
final_local_scores=0.45882744, pem=-1.61336,
similarity=26.55574]
5: 5025 Brisbane_Broncos
[pem_pos=5, final_scores=0.35464174,
final_local_scores=0.2849474, pem=-3.5955102,
similarity=32.146053]
6: 693234 North_Sydney_Bears
[pem_pos=1, final_scores=0.520417,
final_local_scores=0.47525024, pem=-1.09192,
similarity=24.21016]
7: 424878 Sydney_Roosters
[pem_pos=1, final_scores=0.46538234,
final_local_scores=0.40096176, pem=-0.8080674,
similarity=20.449608]
8: 525034 Cronulla-Sutherland_Sharks
[pem_pos=1, final_scores=0.578271,
final_local_scores=0.5295802, pem=-0.7072372,
similarity=23.739508]
9: 454452 Canberra_Raiders
[pem_pos=1, final_scores=0.7161767,
final_local_scores=0.68451667, pem=-2.8732612,
similarity=40.142296]
10: 797018 Newcastle_Knights
[pem_pos=14, final_scores=0.51430804,
final_local_scores=0.4609909, pem=-3.8539813,
similarity=38.80515]
11: 1532351 Western_Suburbs_Magpies
[pem_pos=0, final_scores=0.4038852,
final_local_scores=0.34686804, pem=-0.31132302,
similarity=16.13397]

```
12: 702650 New_Zealand_Warriors
    [pem_pos=13, final_scores=0.5270239,
final_local_scores=0.48998988, pem=-3.4653459,
similarity=37.55708]
13: 1464504 Balmain_Tigers
    [pem_pos=0, final_scores=0.5094665,
final_local_scores=0.46814227, pem=0.0, similarity=18.060013]
14: 869062 Parramatta_Eels
    [pem_pos=1, final_scores=0.7820043,
final_local_scores=0.75236076, pem=-1.3704162,
similarity=33.994194]
15: 516739 Canterbury-Bankstown_Bulldogs
    [pem_pos=0, final_scores=0.6948077,
final_local_scores=0.6532425, pem=0.0, similarity=23.583773]
16: 853881 Illawarra_Steelers
    [pem_pos=3, final_scores=0.4859112,
final_local_scores=0.4535206, pem=-1.9148196,
similarity=28.03676]
17: 2015377 WA_Reds
    [pem_pos=0, final_scores=0.49315673,
final_local_scores=0.48451847, pem=0.0, similarity=18.54871]
18: 471314 Penrith_Panthers
    [pem_pos=2, final_scores=0.6301855,
final_local_scores=0.5854744, pem=-1.61939,
similarity=30.36793]
19: 995013 North_Queensland_Cowboys
    [pem_pos=1, final_scores=0.621389,
final_local_scores=0.58156043, pem=-1.0438483,
similarity=27.121244]
20: 1613020 Gold_Coast_Titans
    [pem_pos=15, final_scores=0.355798,
final_local_scores=0.29821634, pem=-3.7605772,
similarity=33.439686]
21: 1079930 South_Sydney_Rabbitohs
    [pem_pos=0, final_scores=0.885423,
final_local_scores=0.860898, pem=0.0, similarity=29.780626]
22: 2444787 South_Queensland_Crushers
    [pem_pos=0, final_scores=0.5108939,
final_local_scores=0.513504, pem=0.0, similarity=19.413698]
1: 25735 Rugby_league
    [pem_pos=0, final_scores=0.08399956,
final_local_scores=0.033355594, pem=0.0, similarity=5.08511]
2: 853089 St._George_Dragons
    [pem_pos=2, final_scores=0.3669625,
final_local_scores=0.30818453, pem=-2.12389,
similarity=24.836597]
```

Sample: b'372_RUGBY'

```
[1 1 RUGBY UNION 1] - [2 2 All Blacks 2] relive triumph . [3
3 PRETORIA 3] , Aug 25 - Captain [4 4 Sean Fitzpatrick 4] and
his [5 5 All Blacks 5] revisited the test venue today to
relive some of the magic moments of yesterday 's momentous
```

rugby victory over [6 6 South Africa 6] , [7 7 NZPA 7] reported . Most of the test 15 who beat the [8 8 Springboks 8] 33-26 to secure [9 9 New Zealand 9] 's first-ever rugby series in [10 10 South Africa 10] stood in the middle of the empty 50,000-seat [11 11 Loftus Versfeld 11] . Magnificent , ' ' said [12 12 Fitzpatrick 12] , [13 13 New Zealand 13] 's most capped player and the world 's most capped forward . The players relived the moves and tries , the tackles and what might have been as the emotions of victory continued . [14 14 Zinzan Brooke 14] , the only No 8 in test rugby to have scored a dropped goal when he kicked a three-pointer against [15 15 England 15] during last year 's 1 World Cup , added a second to his name yesterday . I was right here , ' ' he said standing at the spot where he had received the ball for the kick . The maul was there and I was going to go in but I thought I should hold off because we had the ball . When (halfback) [16 16 Justin Marshall 16] got the ball he was going to go on the openside where Jon Preston was so I emptied my lung at him to get the ball this way . I just hit through and I was punching the air before the ball got there . It cost me a few bucks at the bar . ' ' The decision to attempt a dropped goal was a spontaneous one , [17 17 Broo ke 17] said . It was just like the 2 World Cup , the ball came and the chance was there . ' ' Centre [18 18 Frank Bunce 18] said he had never felt so exhausted during a match . We were gutted and there was nowhere to hide , they just kept coming at you , ' ' he said . I was gone in the first 20 minutes , completely exhausted , but you had no choice . There was just so much riding on it . It 's amazing just how big this ground was yesterday . ' ' Two-try winger [19 19 Jeff Wilson 19] said he was so tired that he kept asking [20 20 Bu nce 20] where he should be while defending . He told me I 'm buggered too so just hang in there ' , ' ' [21 21 Wilson 21] recalled . About 4000 [22 22 New Zealander 22] supporters were partying into the early hours of today in the [23 23 Sou th African 23] capital . Messages of goodwill continued to roll into the team hotel . [24 24 All Blacks 24] coach [25 25 John Hart 25] said Prime Minister [26 26 Jim Bolger 26] rang him today to offer his congratulations . He thanked us on behalf of the country , which is really nice for the team , and I understand we had tremendous support at home . ' '

- 1: 25405 Rugby_union
- 2: 154957 New_Zealand_national_rugby_union_team
- 3: 25002 Pretoria
- 4: 1444074 Sean_Fitzpatrick
- 5: 154957 New_Zealand_national_rugby_union_team
- 6: 376115 South_Africa_national_rugby_union_team
- 7: 3185516 New_Zealand_Press_Association
- 8: 376115 South_Africa_national_rugby_union_team
- 9: 154957 New_Zealand_national_rugby_union_team
- 10: 17416221 South_Africa
- 11: 2498157 Loftus_Versfeld_Stadium

12: 1444074 Sean_Fitzpatrick
13: 4913064 New_Zealand
14: 740724 Zinzan_Brooke
15: 691024 England_national_rugby_union_team
16: 1809625 Justin_Marshall
17: 740724 Zinzan_Brooke
18: 6093538 Frank_Bunce
19: 905152 Jeff_Wilson_(sportsman)
20: 6093538 Frank_Bunce
21: 905152 Jeff_Wilson_(sportsman)
22: 1286042 New_Zelanders
23: 17416221 South_Africa
24: 154957 New_Zealand_national_rugby_union_team
25: 4996900 John_Hart_(rugby_coach)
26: 362227 Jim_Bolger

1: 526056 1995_Rugby_World_Cup
[winner: 7239 Cricket_World_Cup, pem_pos=4,
final_scores=-0.033568755, final_local_scores=-0.034452558,
pem=-3.1731894, similarity=20.317877]
RECALL MISS: 526056 1995_Rugby_World_Cup]

2: 526056 1995_Rugby_World_Cup
[winner: 7239 Cricket_World_Cup, pem_pos=4,
final_scores=-0.1256567, final_local_scores=-0.13504463,
pem=-3.1731894, similarity=17.316008]
RECALL MISS: 526056 1995_Rugby_World_Cup]

1: 25405 Rugby_union
[pem_pos=0, final_scores=0.45353144,
final_local_scores=0.45049238, pem=0.0, similarity=17.533304]

2: 154957 New_Zealand_national_rugby_union_team
[pem_pos=0, final_scores=0.48040056,
final_local_scores=0.42277235, pem=0.0, similarity=16.706083]

3: 25002 Pretoria
[pem_pos=0, final_scores=0.49027246,
final_local_scores=0.518709, pem=0.0, similarity=19.569025]

4: 1444074 Sean_Fitzpatrick
[pem_pos=0, final_scores=0.5559966,
final_local_scores=0.5621897, pem=0.0, similarity=20.866575]

5: 154957 New_Zealand_national_rugby_union_team
[pem_pos=0, final_scores=0.50553656,
final_local_scores=0.4502296, pem=0.0, similarity=17.525461]

6: 376115 South_Africa_national_rugby_union_team
[pem_pos=3, final_scores=0.42322928,
final_local_scores=0.37788737, pem=-3.598493,
similarity=34.93579]

7: 3185516 New_Zealand_Press_Association
[pem_pos=0, final_scores=0.40514326,
final_local_scores=0.5393403, pem=0.0, similarity=20.184704]

8: 376115 South_Africa_national_rugby_union_team
[pem_pos=0, final_scores=0.48358387,
final_local_scores=0.4438156, pem=0.0, similarity=17.334055]

9: 154957 New_Zealand_national_rugby_union_team
[pem_pos=4, final_scores=0.661082,

```
final_local_scores=0.62013936, pem=-3.8192244,
similarity=43.365444]
10: 17416221 South_Africa
    [pem_pos=0, final_scores=0.336371,
final_local_scores=0.36533117, pem=-0.26251382,
similarity=16.419516]
11: 2498157 Loftus_Versfeld_Stadium
    [pem_pos=0, final_scores=0.48919362,
final_local_scores=0.52562904, pem=0.0, similarity=19.775532]
12: 1444074 Sean_Fitzpatrick
    [pem_pos=0, final_scores=1.0262408,
final_local_scores=1.0830293, pem=0.0, similarity=36.409466]
13: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.51821345,
final_local_scores=0.54726326, pem=-0.23795462,
similarity=21.715174]
14: 740724 Zinzan_Brooke
    [pem_pos=0, final_scores=0.8348004,
final_local_scores=0.8932186, pem=0.0, similarity=30.745138]
15: 691024 England_national_rugby_union_team
    [pem_pos=3, final_scores=0.31035137,
final_local_scores=0.26535743, pem=-3.7874734,
similarity=32.605377]
16: 1809625 Justin_Marshall
    [pem_pos=0, final_scores=0.58693594,
final_local_scores=0.6247456, pem=0.0, similarity=22.733368]
17: 740724 Zinzan_Brooke
    [pem_pos=0, final_scores=0.48571837,
final_local_scores=0.51189953, pem=0.0, similarity=19.365816]
18: 6093538 Frank_Bunce
    [pem_pos=0, final_scores=0.5880983,
final_local_scores=0.6012456, pem=0.0, similarity=22.03208]
19: 905152 Jeff_Wilson_(sportsman)
    [pem_pos=0, final_scores=0.6578904,
final_local_scores=0.67270315, pem=0.0, similarity=24.164516]
20: 6093538 Frank_Bunce
    [pem_pos=0, final_scores=0.69139344,
final_local_scores=0.7215432, pem=0.0, similarity=25.622002]
21: 905152 Jeff_Wilson_(sportsman)
    [pem_pos=0, final_scores=0.6214004,
final_local_scores=0.6328434, pem=0.0, similarity=22.975021]
22: 1286042 New_Zelanders
    [pem_pos=2, final_scores=0.2356296,
final_local_scores=0.2481178, pem=-1.4949476,
similarity=19.6238]
23: 17416221 South_Africa
    [pem_pos=0, final_scores=0.37761283,
final_local_scores=0.41038162, pem=-0.1389451,
similarity=17.091923]
24: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=0, final_scores=0.54880685,
final_local_scores=0.49749577, pem=0.0, similarity=18.93598]
```

```
25: 4996900 John_Hart_(rugby_coach)
    [pem_pos=3, final_scores=0.51782984,
final_local_scores=0.5469925, pem=-1.9445283,
similarity=30.987709]
26: 362227 Jim_Bolger
    [pem_pos=0, final_scores=0.40469962,
final_local_scores=0.50229263, pem=0.0, similarity=19.079126]
```

Sample: b'373_Australian'

```
[1 1 Australian Rules 1] - [2 2 AFL 2] results and
standings . [3 3 MELBOURNE 3] 1996-08-26 Results of [4 4 Australian Rules 4] matches played at the weekend . Played
Sunday : [5 5 Adelaide 5] 14.12 ( 96 ) [6 6 Collingwood 6]
24 . 9 ( 153 ) [7 7 West Coast 7] 24 . 7 ( 151 ) [8 8 Melbourne 8] 11.12 ( 78 ) [9 9 Richmond 9] 28.19 ( 187 ) [10 10 Fitzroy 10] 5 . 6 ( 36 ) Played Saturday : [11 11 Carlton 11]
13.18 ( 96 ) [12 12 Footscray 12] 9.12 ( 66 ) [13 13 Essendon 13] 14.16 ( 100 ) [14 14 Sydney 14] 12.10 ( 82 ) [15 15 St Kilda 15] 9 . 9 ( 63 ) [16 16 Hawthorn 16] 12 . 8 ( 80 ) [17 17 Brisbane 17] 10.11 ( 71 ) [18 18 Fremantle 18] 10.10 ( 70 )
Played Friday : [19 19 North Melbourne 19] 14.12 ( 96 ) [20 20 Geelong 20] 16.13 ( 109 ) Standings ( tabulate under played ,
won , drawn , lost , points for , against , percentage ,
total points ) : [21 21 Brisbane 21] 21 15 1 5 2123 1631
130.2 62 [22 22 Sydney 22] 21 15 1 5 2067 1687 122.5 62 [23 23 West Coast 23] 21 15 0 6 2151 1673 128.6 60 [24 24 North Melbourne 24] 21 15 0 6 2385 1873 127.3 60 [25 25 Carlton 25] 21
14 0 7 2009 1844 108.9 56 [26 26 Geelong 26] 21 13 1 7 2288
1940 117.9 54 [27 27 Essendon 27] 21 13 1 7 2130 1947 109.4
54 [28 28 Richmond 28] 21 11 0 10 2173 1803 120.5 44 [29 29 Hawthorn 29] 21 10 1 10 1791 1820 98.4 42 [30 30 St Kilda 30]
21 9 0 12 1909 1958 97.5 36 [31 31 Collingwood 31] 21 8 0 13
2103 2091 100.6 32 [32 32 Adelaide 32] 21 8 0 13 2158 2183
98.9 32 [33 33 Melbourne 33] 21 7 0 14 1642 2361 69.5 28 [34 3
4 Fremantle 34] 21 6 0 15 1673 1912 87.5 24 [35 35 Footscray 3
5] 21 5 1 15 1578 2060 76.6 22 [36 36 Fitzroy 36] 21 1 0 20
1381 2778 49.7 4 -- Sydney Newsroom 61-2 9373-1800
1: 2403 Australian_rules_football
2: 4732785 Australian_Football_League
3: 17306237 Melbourne
4: 2403 Australian_rules_football
5: 100978 Adelaide_Football_Club
6: 90407 Collingwood_Football_Club
7: 295230 West_Coast_Eagles
8: 97831 Melbourne_Football_Club
9: 100040 Richmond_Football_Club
10: 210973 Fitzroy_Football_Club
11: 89853 Carlton_Football_Club
12: 34094 Western_Bulldogs
13: 10257 Essendon_Football_Club
14: 29210 Sydney_Swans
15: 92922 St_Kilda_Football_Club
```

```
16: 12941233 Hawthorn_Football_Club
17: 5026 Brisbane_Lions
18: 312661 Fremantle_Football_Club
19: 22106 North_Melbourne_Football_Club
20: 13007 Geelong_Football_Club
21: 5026 Brisbane_Lions
22: 29210 Sydney_Swans
23: 295230 West_Coast_Eagles
24: 22106 North_Melbourne_Football_Club
25: 89853 Carlton_Football_Club
26: 13007 Geelong_Football_Club
27: 10257 Essendon_Football_Club
28: 100040 Richmond_Football_Club
29: 12941233 Hawthorn_Football_Club
30: 92922 St_Kilda_Football_Club
31: 90407 Collingwood_Football_Club
32: 100978 Adelaide_Football_Club
33: 97831 Melbourne_Football_Club
34: 312661 Fremantle_Football_Club
35: 34094 Western_Bulldogs
36: 210973 Fitzroy_Football_Club
1: 2403 Australian_rules_football
    [pem_pos=0, final_scores=0.46548104,
final_local_scores=0.4474312, pem=0.0, similarity=17.441952]
2: 4732785 Australian_Football_League
    [pem_pos=0, final_scores=0.4945255,
final_local_scores=0.4481532, pem=-0.73298496,
similarity=21.449585]
3: 17306237 Melbourne
    [pem_pos=0, final_scores=0.42838782,
final_local_scores=0.47514737, pem=-0.16841523,
similarity=19.184927]
4: 2403 Australian_rules_football
    [pem_pos=0, final_scores=0.4825709,
final_local_scores=0.46609926, pem=0.0, similarity=17.999044]
5: 100978 Adelaide_Football_Club
    [pem_pos=1, final_scores=0.60177356,
final_local_scores=0.55662817, pem=-2.7790864,
similarity=35.81371]
6: 90407 Collingwood_Football_Club
    [pem_pos=0, final_scores=0.60551935,
final_local_scores=0.5478763, pem=-0.42581904,
similarity=22.755104]
7: 295230 West_Coast_Eagles
    [pem_pos=2, final_scores=0.652921,
final_local_scores=0.6128932, pem=-1.9809854,
similarity=33.152576]
8: 97831 Melbourne_Football_Club
    [pem_pos=1, final_scores=0.68369794,
final_local_scores=0.6327451, pem=-2.786373,
similarity=38.124817]
9: 100040 Richmond_Football_Club
```

```
[pem_pos=2, final_scores=0.8404961,
final_local_scores=0.80604196, pem=-2.6048384,
similarity=42.309128]
10: 210973 Fitzroy_Football_Club
    [pem_pos=0, final_scores=0.9563659,
final_local_scores=0.9434183, pem=-0.51888555,
similarity=35.06498]
11: 89853 Carlton_Football_Club
    [pem_pos=0, final_scores=0.8469117,
final_local_scores=0.83254343, pem=-0.727443,
similarity=32.89042]
12: 34094 Western_Bulldogs
    [pem_pos=0, final_scores=0.64008623,
final_local_scores=0.5899005, pem=-0.42127568,
similarity=23.984484]
13: 10257 Essendon_Football_Club
    [pem_pos=0, final_scores=0.5894,
final_local_scores=0.531222, pem=-0.09006156,
similarity=20.432205]
14: 29210 Sydney_Swans
    [pem_pos=1, final_scores=0.5695736,
final_local_scores=0.5104839, pem=-4.0017147,
similarity=41.085518]
15: 92922 St_Kilda_Football_Club
    [pem_pos=0, final_scores=0.7556563,
final_local_scores=0.71379304, pem=-0.48976773,
similarity=28.054153]
16: 12941233 Hawthorn_Football_Club
    [pem_pos=0, final_scores=0.7307991,
final_local_scores=0.6875568, pem=-0.5078096,
similarity=27.369326]
17: 5026 Brisbane_Lions
    [pem_pos=2, final_scores=0.33543754,
final_local_scores=0.2711038, pem=-3.4494781,
similarity=30.938791]
18: 312661 Fremantle_Football_Club
    [pem_pos=1, final_scores=0.6218572,
final_local_scores=0.57925224, pem=-1.2687272,
similarity=28.275291]
19: 22106 North_Melbourne_Football_Club
    [pem_pos=0, final_scores=0.7057052,
final_local_scores=0.65817153, pem=0.0, similarity=23.730864]
20: 13007 Geelong_Football_Club
    [pem_pos=1, final_scores=0.47009206,
final_local_scores=0.40278035, pem=-0.8939384,
similarity=20.970858]
21: 5026 Brisbane_Lions
    [pem_pos=2, final_scores=0.335595,
final_local_scores=0.27127582, pem=-3.4494781,
similarity=30.943922]
22: 29210 Sydney_Swans
    [pem_pos=1, final_scores=0.6298509,
```

```
final_local_scores=0.57632774, pem=-4.0017147,
similarity=43.050426]
23: 295230 West_Coast_Eagles
    [pem_pos=2, final_scores=0.58914024,
final_local_scores=0.5432224, pem=-1.9809854,
similarity=31.073462]
24: 22106 North_Melbourne_Football_Club
    [pem_pos=0, final_scores=0.68986,
final_local_scores=0.640863, pem=0.0, similarity=23.214344]
25: 89853 Carlton_Football_Club
    [pem_pos=0, final_scores=0.77600724,
final_local_scores=0.755091, pem=-0.727443,
similarity=30.579082]
26: 13007 Geelong_Football_Club
    [pem_pos=1, final_scores=0.41547257,
final_local_scores=0.34311682, pem=-0.8939384,
similarity=19.190378]
27: 10257 Essendon_Football_Club
    [pem_pos=0, final_scores=0.52747726,
final_local_scores=0.46358085, pem=-0.09006156,
similarity=18.41366]
28: 100040 Richmond_Football_Club
    [pem_pos=2, final_scores=0.7169091,
final_local_scores=0.67104197, pem=-2.6048384,
similarity=38.28046]
29: 12941233 Hawthorn_Football_Club
    [pem_pos=0, final_scores=0.75559455,
final_local_scores=0.71464205, pem=-0.5078096,
similarity=28.177605]
30: 92922 St_Kilda_Football_Club
    [pem_pos=0, final_scores=0.6316976,
final_local_scores=0.57838696, pem=-0.48976773,
similarity=24.013367]
31: 90407 Collingwood_Football_Club
    [pem_pos=0, final_scores=0.5733296,
final_local_scores=0.51271397, pem=-0.42581904,
similarity=21.70579]
32: 100978 Adelaide_Football_Club
    [pem_pos=1, final_scores=0.56833214,
final_local_scores=0.5200985, pem=-2.7790864,
similarity=34.72359]
33: 97831 Melbourne_Football_Club
    [pem_pos=1, final_scores=0.58884424,
final_local_scores=0.5291318, pem=-2.786373,
similarity=35.03279]
34: 312661 Fremantle_Football_Club
    [pem_pos=1, final_scores=0.5111617,
final_local_scores=0.4583342, pem=-1.2687272,
similarity=24.666857]
35: 34094 Western_Bulldogs
    [pem_pos=0, final_scores=0.65193826,
final_local_scores=0.602847, pem=-0.42127568,
```

```
similarity=24.370832]
36: 210973 Fitzroy_Football_Club
    [pem_pos=0, final_scores=0.77355164,
final_local_scores=0.7437214, pem=-0.51888555,
similarity=29.105623]

Sample: b'374_Rugby'
Rugby league-Australian rugby league results . [1 1 SYDNEY 1]
1996-08-26 Results of [2 2 Australian 2] rugby league matches
played at the weekend . Played Sunday : [3 3 Sydney Bulldogs 3
] 17 [4 4 South Queensland 4] 16 [5 5 Brisbane 5] 38 [6 6 Gold
Coast 6] 10 [7 7 North Sydney 7] 46 [8 8 South Sydney 8] 4 [9
9 Illawarra 9] 42 [10 10 Penrith 10] 2 [11 11 St George 11]
20 [12 12 North Queensland 12] 24 [13 13 Manly 13] 42 [14 14 W
estern Suburbs 14] 12 Played Saturday : [15 15 Parramatta 15]
14 [16 16 Sydney Tigers 16] 26 [17 17 Newcastle 17] 24 [18 18
Western Reds 18] 20 Played Friday : [19 19 Canberra 19] 30 [20
20 Auckland 20] 6 -- Sydney Newsroom 61-2 9373-1800
1: 27862 Sydney
2: 4689264 Australia
3: 516739 Canterbury-Bankstown_Bulldogs
4: 2444787 South_Queensland_Crushers
5: 5025 Brisbane_Broncos
6: 1613020 Gold_Coast_Titans
7: 693234 North_Sydney_Bears
8: 1079930 South_Sydney_Rabbitohs
9: 853881 Illawarra_Steelers
10: 471314 Penrith_Panthers
11: 853089 St._George_Dragons
12: 995013 North_Queensland_Cowboys
13: 885301 Manly-Warringah_Sea_Eagles
14: 1532351 Western_Suburbs_Magpies
15: 869062 Parramatta_Eels
16: 1464504 Balmain_Tigers
17: 797018 Newcastle_Knights
18: 2015377 WA_Reds
19: 454452 Canberra_Raiders
20: 702650 New_Zealand_Warriors
1: 27862 Sydney
    [pem_pos=0, final_scores=0.54661316,
final_local_scores=0.61625904, pem=-0.1443489,
similarity=23.265104]
2: 4689264 Australia
    [pem_pos=0, final_scores=0.6158205,
final_local_scores=0.7151548, pem=-0.491009,
similarity=28.101542]
3: 516739 Canterbury-Bankstown_Bulldogs
    [pem_pos=0, final_scores=0.6670139,
final_local_scores=0.62256944, pem=0.0, similarity=22.668427]
4: 2444787 South_Queensland_Crushers
    [pem_pos=0, final_scores=0.5776352,
final_local_scores=0.58304995, pem=0.0, similarity=21.489086]
```

```
5: 5025 Brisbane_Broncos
    [pem_pos=5, final_scores=0.4629236,
final_local_scores=0.40142953, pem=-3.5955102,
similarity=35.622112]
6: 1613020 Gold_Coast_Titans
    [pem_pos=15, final_scores=0.27399057,
final_local_scores=0.20815158, pem=-3.7605772,
similarity=30.751976]
7: 693234 North_Sydney_Bears
    [pem_pos=1, final_scores=0.52979374,
final_local_scores=0.48525953, pem=-1.09192,
similarity=24.508856]
8: 1079930 South_Sydney_Rabbitohs
    [pem_pos=0, final_scores=0.8582354,
final_local_scores=0.8317825, pem=0.0, similarity=28.911762]
9: 853881 Illawarra_Steelers
    [pem_pos=3, final_scores=0.52089125,
final_local_scores=0.49004263, pem=-1.9148196,
similarity=29.126648]
10: 471314 Penrith_Panthers
    [pem_pos=2, final_scores=0.6143108,
final_local_scores=0.56735563, pem=-1.61939,
similarity=29.82723]
11: 853089 St_George_Dragons
    [pem_pos=2, final_scores=0.43070233,
final_local_scores=0.37534255, pem=-2.12389,
similarity=26.840727]
12: 995013 North_Queensland_Cowboys
    [pem_pos=1, final_scores=0.62480956,
final_local_scores=0.58421713, pem=-1.0438483,
similarity=27.200525]
13: 885301 Manly-Warringah_Sea_Eagles
    [pem_pos=1, final_scores=0.52700233,
final_local_scores=0.47086412, pem=-1.61336,
similarity=26.91494]
14: 1532351 Western_Suburbs_Magpies
    [pem_pos=0, final_scores=0.39175826,
final_local_scores=0.33208936, pem=-0.31132302,
similarity=15.692945]
15: 869062 Parramatta_Eels
    [pem_pos=1, final_scores=0.76788074,
final_local_scores=0.7365261, pem=-1.3704162,
similarity=33.521656]
16: 1464504 Balmain_Tigers
    [pem_pos=0, final_scores=0.53669,
final_local_scores=0.49633372, pem=0.0, similarity=18.9013]
17: 797018 Newcastle_Knights
    [pem_pos=14, final_scores=0.5456452,
final_local_scores=0.49439085, pem=-3.8539813,
similarity=39.801872]
18: 2015377 WA_Reds
    [pem_pos=0, final_scores=0.40628666,
```

```
final_local_scores=0.38748395, pem=0.0, similarity=15.6530075]
19: 454452 Canberra_Raiders
    [pem_pos=1, final_scores=0.6587857,
final_local_scores=0.6210005, pem=-2.8732612,
similarity=38.24685]
20: 702650 New_Zealand_Warriors
    [pem_pos=13, final_scores=0.5112097,
final_local_scores=0.47226644, pem=-3.4653459,
similarity=37.028175]
```

Sample: b'375_South'

[1 1 South Africa 1] yet to hear from apartheid enforcers . Anton Ferreira [2 2 CAPE TOWN 2] 1996-08-25 [3 3 South Africa 3] 's truth commission begins issuing subpoenas this week in a bid to dig beneath the political rationales and find the sinister figures who have the blood of the country 's race war on their hands . Leaders of the major parties involved , from right-wing whites to radical blacks , appeared before Archbishop [4 4 Desmond Tutu 4] 's Truth and Reconciliation Commission last week to paint the broad picture of their actions for or against apartheid . Most , including former president [5 5 F.W. de Klerk 5] and [6 6 African National Congress 6] Deputy President [7 7 Thabo Mbeki 7] , offered apologies for any mistakes they had made and accepted broad responsibility for the actions of their foot soldiers . But none named those guilty of ordering or carrying out any of the gross violations of human rights which [8 8 Tutu 8] is investigating . Human rights lawyer [9 9 Brian Currin 9] said the hearings last week were not intended as a form of confessional and that those who personally committed crimes during apartheid would testify only before a separate arm of the commission , the amnesty committee . " I do n't think one should have expected more than what one got , " he said . " That was not the amnesty committee where perpetrators are expected to open their hearts and souls and to tell it all . " The commission , which has the power to grant amnesty to those who confess to abuses , has begun hearing the testimony from people already in jail for their deeds . But others , such as self-confessed secret police hit-squad leader [10 10 Dirk Coetze 10] , have yet to testify . [11 11 Tutu 11] 's deputy chairman , Alex Boraine , told reporters the commission would begin issuing subpoenas to suspects who refused to appear voluntarily some time this week . He added that former hardline apartheid president [12 12 P.W. Botha 12] could be among those called . But [13 13 Currin 13] , who is advising several people regarded as perpetrators , said this was not the best way to achieve the commission 's aims . " A person can be forced to appear , but the only way one is going to get to the truth in its totality is if people feel it is a good idea to go to the amnesty committee . At the moment this is not the case . " He cited [14 14 Coetze 14] , who was charged with murder after confessing in media

interviews to dirty tricks . His trial is due to start in December but the truth commission intends to decide on his amnesty application before that . [15 15 Currin 15] said the law had to be changed so all judicial prosecutions were automatically suspended for those who approached the truth commission . Last week 's submissions to the commission by the [16 16 ANC 16] , de Klerk 's [17 17 National Party 17] and the right-wing [18 18 Freedom Front 18] of General [19 19 Constand Viljoen 19] left many South Africans unsatisfied . Boraine said the picture was incomplete because officers of the apartheid-era security forces had yet to make their scheduled , separate submission , but one black caller to a radio talk show declared : " This whole thing is utterly useless , it does n't help us at all . What the people expect is to have houses , to have jobs . " Political scientist Jannie Gagiano said he doubted that the [20 20 National Party 20] , which implemented apartheid in 1948 and began dismantling it in 1990 , felt it carried a burden of guilt and needed to be exculpated through the commission . " Therefore I have some doubts about achieving some form of reconciliation . One has to feel a bit guilty to feel the need for reconciling yourself to a historical adversary , " he said . Professor Tom Lodge of the [21 21 University of the Witwatersrand 21] demurred , saying : " In the irritation , in the jokes , in the anger that white South Africans express about the commission , I think there 's a moral uneasiness , and I think that 's healthy . Responsibility is percolating downwards . "

1: 17416221 South_Africa
2: 6653 Cape_Town
3: 17416221 South_Africa
4: 8401973 Desmond_Tutu
5: 11479 F._W._de_Klerk
6: 2503 African_National_Congress
7: 31350 Thabo_Mbeki
8: 8401973 Desmond_Tutu
9: 14059389 Brian_Currin
10: 11288911 Dirk_Coetzee
11: 8401973 Desmond_Tutu
12: 383252 P._W._Botha
13: 14059389 Brian_Currin
14: 11288911 Dirk_Coetzee
15: 14059389 Brian_Currin
16: 2503 African_National_Congress
17: 390975 National_Party_(South_Africa)
18: 449582 Freedom_Front_Plus
19: 2030631 Constand_Viljoen
20: 390975 National_Party_(South_Africa)
21: 320467 University_of_the_Witwatersrand
1: 17416221 South_Africa
[pem_pos=0, final_scores=0.42331678,
final_local_scores=0.45960993, pem=-0.26251382,

```
similarity=19.23298]
2: 6653 Cape_Town
    [pem_pos=0, final_scores=0.44480258,
final_local_scores=0.47220057, pem=0.0, similarity=18.18112]
3: 17416221 South_Africa
    [pem_pos=0, final_scores=0.44844973,
final_local_scores=0.48706383, pem=-0.26251382,
similarity=20.05226]
4: 8401973 Desmond_Tutu
    [pem_pos=0, final_scores=0.56844085,
final_local_scores=0.5889815, pem=0.0, similarity=21.666096]
5: 11479 F._W._de_Klerk
    [pem_pos=0, final_scores=0.47392327,
final_local_scores=0.45679408, pem=0.0, similarity=17.72136]
6: 2503 African_National_Congress
    [pem_pos=0, final_scores=0.3953939,
final_local_scores=0.3553776, pem=0.0, similarity=14.69489]
7: 31350 Thabo_Mbeki
    [pem_pos=0, final_scores=0.63317704,
final_local_scores=0.63221985, pem=0.0, similarity=22.956415]
8: 8401973 Desmond_Tutu
    [pem_pos=0, final_scores=0.78210515,
final_local_scores=0.83207244, pem=0.0, similarity=28.920414]
9: 14059389 Brian_Currin
    [pem_pos=0, final_scores=0.56767094,
final_local_scores=0.6382276, pem=0.0, similarity=23.135696]
10: 11288911 Dirk_Coetzee
    [pem_pos=0, final_scores=0.5516328,
final_local_scores=0.6325873, pem=0.0, similarity=22.96738]
11: 8401973 Desmond_Tutu
    [pem_pos=0, final_scores=0.7869745,
final_local_scores=0.8373915, pem=0.0, similarity=29.079145]
12: 383252 P._W._Botha
    [pem_pos=0, final_scores=0.6640912,
final_local_scores=0.6833367, pem=0.0, similarity=24.481842]
13: 14059389 Brian_Currin
    [pem_pos=1, final_scores=0.6924048,
final_local_scores=0.78794, pem=0.0, similarity=27.603415]
14: 11288911 Dirk_Coetzee
    [pem_pos=0, final_scores=0.8241901,
final_local_scores=0.944545, pem=0.0, similarity=32.276817]
15: 14059389 Brian_Currin
    [pem_pos=1, final_scores=0.6581382,
final_local_scores=0.750509, pem=0.0, similarity=26.486399]
16: 2503 African_National_Congress
    [pem_pos=0, final_scores=0.51590526,
final_local_scores=0.48701793, pem=-0.05902751,
similarity=18.9443]
17: 390975 National_Party_(South_Africa)
    [pem_pos=0, final_scores=0.43849045,
final_local_scores=0.42302376, pem=-0.99382544,
similarity=22.118164]
```

```
18: 449582 Freedom_Front_Plus
    [pem_pos=0, final_scores=0.5696456,
final_local_scores=0.64, pem=0.0, similarity=23.18859]
19: 2030631 Constand_Viljoen
    [pem_pos=0, final_scores=0.52787596,
final_local_scores=0.58067286, pem=0.0, similarity=21.41815]
20: 390975 National_Party_(South_Africa)
    [pem_pos=0, final_scores=0.46656066,
final_local_scores=0.45368618, pem=-0.99382544,
similarity=23.033192]
21: 320467 University_of_the_Witwatersrand
    [pem_pos=0, final_scores=0.4167394,
final_local_scores=0.4464401, pem=0.0, similarity=17.412376]
```

Sample: b'376_Moslem'

[1 1 Moslem 1] refugees seek to vote in Serb-held town . Samir Arnaut [2 2 MATUZICI 2] , Bosnia 1996-08-25 Thousands of [3 3 Moslem 3] refugees denounced Bosnia 's elections as a farce on Sunday because [4 4 Serb 4] incomers would be able to vote in their old home town , cementing partition of the country . They said they were ready to force their way back across post-war ethnic lines to Serb-held [5 5 Doboј 5] , one of several towns where [1 1 NATO 1] troops fear violence involving refugees determined to vote where they once lived . The refugees , rallying in [6 6 Matuzici 6] on government territory two km from [7 7 Doboј 7] in northeast Bosnia , waved banners calling the polls a farce staged by the [8 8 United Nations 8] and the [9 9 European Union 9] . " We demand to vote in [10 10 Doboј 10] , " other banners said . Bosnia 's Moslem-led central government and many displaced [11 11 Moslems 11] are angered by a key provision of the Western-organised September 14 elections allowing people to vote in post-war " new places of residence . " As a result , separatist [12 12 Serb 12] authorities have packed former [13 13 Moslem 13] majority towns with refugees of their own or registered other [14 14 Serbs 14] to vote there . Critics say that what was billed as an electoral process to reintegrate Bosnia as a single , multi-ethnic state is shaping up as a referendum on partition , de facto or de jure . " Only our physical presence in [15 15 Doboј 15] will mean that the 2 Dayton peace treaty has truly been implemented , " Edhem Efendija Camdzic , [16 16 Doboј 16] 's [17 17 Islamic 17] imam-in-exile , told the refugees . " The main point of this rally is to highlight to the world powers what misfortune they brought upon us , " said Reuf Mehmedagic , head of [18 18 Doboј 18] municipality-in-exile . " Eleven thousand [19 19 Serbs 19] who came from elsewhere to [20 20 Doboј 20] will vote there . How can we then expect the reintegration of Bosnia ? But no one will stop us from returning to our homes . " Mirhunisa Komarica , a government refugee official , said : " We want to vote where we were thrown out from . This is a protest of warning and the next step is entering our town

using all means possible , so let them shoot . " Bosnian Vice President [21 21 Ejup Ganic 21] , a [22 22 Moslem 22] , told the refugees : " We have a message for the [23 23 Serbs 23] who are now in our homes not to plan the future of their children there because there will be no good fortune in that . " To loud applause , he added : " We will enter [24 24 Doboј 24] , untie the [25 25 Doboј 25] knot and ensure free movement for all . We have to enter [26 26 Doboј 26] to free the [27 27 Serbs 27] from their own (separatist) politics . " The 3 Dayton peace accords guaranteed refugees the right to return to their homes and ensured freedom of movement across ethnic lines . But local police , controlled by nationalist parties in [28 28 Moslem 28] , [29 29 Serb 29] and [30 30 Croat 30] sectors of Bosnia , and civilian mobs have turned ceasefire lines into virtually impassable borders . The majority of refugees from Bosnia 's 1992-95 war are [31 31 Moslems 31] from the Serb-controlled north and east . NATO-led peace troops beefed up their presence in the [32 32 Doboј 32] area on Sunday to deter any sudden emotional attempt by the refugees to cross the Inter-Entity Boundary Line into [33 33 Doboј 33] . But the crowd dispersed without incident . [34 34 Moslems 34] and [35 35 Serbs 35] have scuffled several times along the line in the past when [36 36 Moslem 36] refugees tried to surge into the town .

1: 6037917 Islam
2: 28384119 Matuzići
3: 6037917 Islam
4: 29265 Serbia
5: 294608 Doboј
6: 28384119 Matuzići
7: 294608 Doboј
8: 31769 United_Nations
9: 9317 European_Union
10: 294608 Doboј
11: 6037917 Islam
12: 29265 Serbia
13: 6037917 Islam
14: 59512 Serbs
15: 294608 Doboј
16: 294608 Doboј
17: 6037917 Islam
18: 294608 Doboј
19: 59512 Serbs
20: 294608 Doboј
21: 21386224 Ejup_Ganić
22: 6037917 Islam
23: 59512 Serbs
24: 294608 Doboј
25: 294608 Doboј
26: 294608 Doboј
27: 59512 Serbs
28: 6037917 Islam

29: 29265 Serbia
30: 375905 Independent_State_of_Croatia
31: 6037917 Islam
32: 294608 Doboј
33: 294608 Doboј
34: 6037917 Islam
35: 59512 Serbs
36: 6037917 Islam
1: 21710997 NATO_intervention_in_Bosnia_and_Herzegovina
[winner: 21133 NATO, pem_pos=0, final_scores=0.051272318,
final_local_scores=0.08915217, pem=0.0, similarity=6.7501917 |
gt: 21710997
NATO_intervention_in_Bosnia_and_Herzegovina, pem_pos=6,
final_scores=-0.028163657, final_local_scores=0.06024587,
pem=-2.6388168, similarity=20.237865]
2: 168269 Dayton_Agreement
[winner: 8253 Dayton,_Ohio, pem_pos=0,
final_scores=-0.24941427, final_local_scores=-0.20713188,
pem=-0.9297534, similarity=2.964628 |
RECALL MISS: 168269 Dayton_Agreement]
3: 168269 Dayton_Agreement
[winner: 8253 Dayton,_Ohio, pem_pos=0,
final_scores=-0.12818748, final_local_scores=-0.07471003,
pem=-0.9297534, similarity=6.9163604 |
RECALL MISS: 168269 Dayton_Agreement]
1: 6037917 Islam
[pem_pos=1, final_scores=0.42021745,
final_local_scores=0.4356866, pem=-1.2150909,
similarity=23.699322]
2: 28384119 Matuzići
[pem_pos=0, final_scores=0.3557093,
final_local_scores=0.5642317, pem=0.0, similarity=20.927511]
3: 6037917 Islam
[pem_pos=1, final_scores=0.46519464,
final_local_scores=0.4848174, pem=-1.2150909,
similarity=25.165483]
4: 29265 Serbia
[pem_pos=1, final_scores=0.584106,
final_local_scores=0.5700499, pem=-2.1726613,
similarity=32.91641]
5: 294608 Doboј
[pem_pos=0, final_scores=0.5810901,
final_local_scores=0.5627567, pem=0.0, similarity=20.883495]
6: 28384119 Matuzići
[pem_pos=0, final_scores=0.4415264,
final_local_scores=0.6579738, pem=0.0, similarity=23.724964]
7: 294608 Doboј
[pem_pos=0, final_scores=0.50743705,
final_local_scores=0.48230195, pem=0.0, similarity=18.482565]
8: 31769 United_Nations
[pem_pos=0, final_scores=0.38285714,
final_local_scores=0.44233203, pem=0.0, similarity=17.289783]

```
9: 9317 European_Union
    [pem_pos=0, final_scores=0.43364358,
final_local_scores=0.48385924, pem=-0.005814796,
similarity=18.56066]
10: 294608 Doboj
    [pem_pos=0, final_scores=0.58004427,
final_local_scores=0.56161433, pem=0.0, similarity=20.849405]
11: 6037917 Islam
    [pem_pos=0, final_scores=0.58252114,
final_local_scores=0.6129787, pem=0.0, similarity=22.38222]
12: 29265 Serbia
    [pem_pos=1, final_scores=0.6361062,
final_local_scores=0.62685215, pem=-2.1726613,
similarity=34.6115]
13: 6037917 Islam
    [pem_pos=1, final_scores=0.44892794,
final_local_scores=0.46704847, pem=-1.2150909,
similarity=24.635223]
14: 59512 Serbs
    [pem_pos=0, final_scores=0.45032907,
final_local_scores=0.40364337, pem=-0.02793329,
similarity=16.287142]
15: 294608 Doboj
    [pem_pos=0, final_scores=0.6017323,
final_local_scores=0.5853052, pem=0.0, similarity=21.556389]
16: 294608 Doboj
    [pem_pos=0, final_scores=0.48597914,
final_local_scores=0.45886248, pem=0.0, similarity=17.783085]
17: 6037917 Islam
    [pem_pos=0, final_scores=0.44380254,
final_local_scores=0.461444968, pem=-0.21141872,
similarity=19.01002]
18: 294608 Doboj
    [pem_pos=0, final_scores=0.49485737,
final_local_scores=0.46856058, pem=0.0, similarity=18.072495]
19: 59512 Serbs
    [pem_pos=0, final_scores=0.4372763,
final_local_scores=0.38938522, pem=-0.02793329,
similarity=15.86165]
20: 294608 Doboj
    [pem_pos=0, final_scores=0.6121833,
final_local_scores=0.5967213, pem=0.0, similarity=21.897066]
21: 21386224 Ejup_Ganić
    [pem_pos=0, final_scores=0.46169573,
final_local_scores=0.5696154, pem=0.0, similarity=21.088173]
22: 6037917 Islam
    [pem_pos=1, final_scores=0.39049345,
final_local_scores=0.4032176, pem=-1.2150909,
similarity=22.730383]
23: 59512 Serbs
    [pem_pos=0, final_scores=0.41233253,
final_local_scores=0.36213794, pem=-0.02793329,
```

```
similarity=15.048537]
24: 294608 Doboj
    [pem_pos=0, final_scores=0.614563,
final_local_scores=0.59932077, pem=0.0, similarity=21.97464]
25: 294608 Doboj
    [pem_pos=0, final_scores=0.58535504,
final_local_scores=0.56741554, pem=0.0, similarity=21.022524]
26: 294608 Doboj
    [pem_pos=0, final_scores=0.53416765,
final_local_scores=0.5115011, pem=0.0, similarity=19.353926]
27: 59512 Serbs
    [pem_pos=0, final_scores=0.42433506,
final_local_scores=0.3752489, pem=-0.02793329,
similarity=15.4397955]
28: 6037917 Islam
    [pem_pos=1, final_scores=0.41211957,
final_local_scores=0.4268409, pem=-1.2150909,
similarity=23.435349]
29: 29265 Serbia
    [pem_pos=1, final_scores=0.5374833,
final_local_scores=0.5191215, pem=-2.1726613,
similarity=31.396605]
30: 375905 Independent_State_of_Croatia
    [pem_pos=2, final_scores=0.121291175,
final_local_scores=0.10256392, pem=-1.5931925,
similarity=15.814453]
31: 6037917 Islam
    [pem_pos=0, final_scores=0.5652203,
final_local_scores=0.59408015, pem=0.0, similarity=21.818249]
32: 294608 Doboj
    [pem_pos=0, final_scores=0.5062698,
final_local_scores=0.48102695, pem=0.0, similarity=18.444517]
33: 294608 Doboj
    [pem_pos=0, final_scores=0.5381881,
final_local_scores=0.5158928, pem=0.0, similarity=19.484983]
34: 6037917 Islam
    [pem_pos=0, final_scores=0.51663977,
final_local_scores=0.54101336, pem=0.0, similarity=20.23463]
35: 59512 Serbs
    [pem_pos=0, final_scores=0.33681738,
final_local_scores=0.2796491, pem=-0.02793329,
similarity=12.5869055]
36: 6037917 Islam
    [pem_pos=1, final_scores=0.4948576,
final_local_scores=0.51721966, pem=-1.2150909,
similarity=26.132431]
1: 21133 NATO
    [pem_pos=0, final_scores=0.051272318,
final_local_scores=0.08915217, pem=0.0, similarity=6.7501917]
```

Sample: b'377_Eight'

Eight killed in [1 1 Moscow 1] casino blaze . [2 2 MOSCOW 2]

1996-08-25 Eight people died on Sunday in a blaze at a [3 3 Moscow 3] casino which the fire service said might have been started deliberately , [4 4 Interfax 4] news agency said . The number of casinos has soared in [5 5 Moscow 5] since the collapse of communism . The mayor has said he wants to cut their number to five as part of a war against organised crime . President [6 6 Boris Yeltsin 6] signed a decree on fighting crime in July and handed wide-ranging powers to security chief [7 7 Alexander Lebed 7] , currently engaged in making peace in breakaway [8 8 Chechnya 8] .

1: 19004 Moscow
2: 19004 Moscow
3: 19004 Moscow
4: 4921404 Interfax
5: 19004 Moscow
6: 23278520 Boris_Yeltsin
7: 431579 Alexander_Lebed
8: 6095 Chechnya
1: 19004 Moscow
[pem_pos=0, final_scores=0.59518963,
final_local_scores=0.55424964, pem=-0.106582455,
similarity=21.20924]
2: 19004 Moscow
[pem_pos=0, final_scores=0.73720664,
final_local_scores=0.7093817, pem=-0.106582455,
similarity=25.83869]
3: 19004 Moscow
[pem_pos=0, final_scores=0.58379984,
final_local_scores=0.54180807, pem=-0.106582455,
similarity=20.83796]
4: 4921404 Interfax
[pem_pos=0, final_scores=0.675468,
final_local_scores=0.74222076, pem=0.0, similarity=26.239061]
5: 19004 Moscow
[pem_pos=0, final_scores=0.4878711,
final_local_scores=0.43702048, pem=-0.106582455,
similarity=17.710888]
6: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.7018301,
final_local_scores=0.71907234, pem=0.0, similarity=25.548265]
7: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.76997584,
final_local_scores=0.83416337, pem=0.0, similarity=28.982811]
8: 6095 Chechnya
[pem_pos=0, final_scores=0.55414087,
final_local_scores=0.571281, pem=0.0, similarity=21.137878]

Sample: b'378_Russian'

[1 1 Russian 1] troops start pullout , but not in [2 2 Grozny 2] . [3 3 SHATOI 3] , [4 4 Russia 4] 1996-08-25 [5 5 Russian 5] troops began to pull out from southern [6 6 Chechnya 6] on Sunday under a ceasefire agreement between [7 7 Russian 7]

security chief [8 8 Alexander Lebed 8] and rebel leaders . But in the capital [9 9 Grozny 9] , the commander of [10 10 Russian 10] Interior Ministry forces in [11 11 Chechnya 11] , General Anatoly Shkirkko , told [12 12 Interfax 12] news agency he was delaying a pullout of troops there after a group of [13 13 Chechens 13] disarmed an armoured column . [14 14 Reuters 14] cameraman Liutauras Stremaitis said a column of around 40 vehicles , including tanks , armoured personnel carriers , artillery cannons and lorries , escorted by [15 15 Chechen 15] rebels , pulled out of the village of [16 16 Shatoy 16] towards the border , around 50 km (31 miles) to the north . In [17 17 Grozny 17] , Shkirkko told [18 18 Interfax 18] he was suspending the pullout of troops from the capital until weapons seized by the [19 19 Chechens 19] were returned . [20 20 Chechen 20] rebel spokesman [21 21 Movladi U dugov 21] confirmed the weapons had been seized but that it was a maverick group of [22 22 Chechens 22] . He said later that the rebels had handed them over . The pullout of the [23 23 Russian 23] troops is a key element of the peace plan brokered by [24 24 Lebed 24] , which aims to end the 20-month [25 25 Chechnya 25] conflict .

1: 25391 Russia
2: 140790 Grozny
3: 4142015 Shatoy
4: 25391 Russia
5: 25391 Russia
6: 6095 Chechnya
7: 25391 Russia
8: 431579 Alexander_Lebed
9: 140790 Grozny
10: 25391 Russia
11: 6095 Chechnya
12: 4921404 Interfax
13: 6095 Chechnya
14: 18998750 Reuters
15: 6095 Chechnya
16: 4142015 Shatoy
17: 140790 Grozny
18: 4921404 Interfax
19: 6095 Chechnya
20: 6095 Chechnya
21: 6054747 Movladi_Udugov
22: 6095 Chechnya
23: 25391 Russia
24: 431579 Alexander_Lebed
25: 6095 Chechnya
1: 25391 Russia
[pem_pos=1, final_scores=0.9826649,
final_local_scores=1.0150976, pem=-1.0515022,
similarity=40.100483]
2: 140790 Grozny
[pem_pos=0, final_scores=0.56611365,

```
final_local_scores=0.55068654, pem=0.0, similarity=20.523298]
3: 4142015 Shatoy
    [pem_pos=0, final_scores=0.38125104,
final_local_scores=0.4952309, pem=0.0, similarity=18.868391]
4: 25391 Russia
    [pem_pos=0, final_scores=0.84447914,
final_local_scores=0.8641507, pem=-0.20299491,
similarity=30.98161]
5: 25391 Russia
    [pem_pos=1, final_scores=0.88449174,
final_local_scores=0.90785843, pem=-1.0515022,
similarity=36.900253]
6: 6095 Chechnya
    [pem_pos=0, final_scores=0.58446246,
final_local_scores=0.5382635, pem=0.0, similarity=20.152569]
7: 25391 Russia
    [pem_pos=1, final_scores=0.8763526,
final_local_scores=0.8989677, pem=-1.0515022,
similarity=36.634937]
8: 431579 Alexander_Lebed
    [pem_pos=0, final_scores=0.7914061,
final_local_scores=0.8391717, pem=0.0, similarity=29.13227]
9: 140790 Grozny
    [pem_pos=0, final_scores=0.62025994,
final_local_scores=0.6098332, pem=0.0, similarity=22.288351]
10: 25391 Russia
    [pem_pos=1, final_scores=0.79508466,
final_local_scores=0.81019473, pem=-1.0515022,
similarity=33.98577]
11: 6095 Chechnya
    [pem_pos=0, final_scores=0.59382504,
final_local_scores=0.54849064, pem=0.0, similarity=20.457767]
12: 4921404 Interfax
    [pem_pos=0, final_scores=0.65483814,
final_local_scores=0.7033369, pem=0.0, similarity=25.07869]
13: 6095 Chechnya
    [pem_pos=2, final_scores=0.5612175,
final_local_scores=0.51287186, pem=-0.96860826,
similarity=24.662273]
14: 18998750 Reuters
    [pem_pos=0, final_scores=0.33599907,
final_local_scores=0.37821978, pem=0.0, similarity=15.376545]
15: 6095 Chechnya
    [pem_pos=2, final_scores=0.5529633,
final_local_scores=0.50385547, pem=-0.8617233,
similarity=23.81195]
16: 4142015 Shatoy
    [pem_pos=0, final_scores=0.47783697,
final_local_scores=0.6007364, pem=0.0, similarity=22.016884]
17: 140790 Grozny
    [pem_pos=0, final_scores=0.5775458,
final_local_scores=0.5631744, pem=0.0, similarity=20.895962]
```

```
18: 4921404 Interfax
    [pem_pos=0, final_scores=0.46542615,
final_local_scores=0.49643308, pem=0.0, similarity=18.904266]
19: 6095 Chechnya
    [pem_pos=2, final_scores=0.56199616,
final_local_scores=0.5137225, pem=-0.96860826,
similarity=24.687658]
20: 6095 Chechnya
    [pem_pos=2, final_scores=0.51257896,
final_local_scores=0.45974165, pem=-0.8617233,
similarity=22.495506]
21: 6054747 Movladi_Udugov
    [pem_pos=0, final_scores=0.44833505,
final_local_scores=0.4902671, pem=0.0, similarity=18.72026]
22: 6095 Chechnya
    [pem_pos=2, final_scores=0.5420383,
final_local_scores=0.49192154, pem=-0.96860826,
similarity=24.037075]
23: 25391 Russia
    [pem_pos=1, final_scores=0.90677685,
final_local_scores=0.93220145, pem=-1.0515022,
similarity=37.626698]
24: 431579 Alexander_Lebed
    [pem_pos=0, final_scores=0.64646214,
final_local_scores=0.69005495, pem=0.0, similarity=24.68233]
25: 6095 Chechnya
    [pem_pos=0, final_scores=0.6183943,
final_local_scores=0.5753289, pem=0.0, similarity=21.258675]
```

Sample: b'379_Leftist'

Leftist [1 1 Mexican 1] armed group says troops in capital . [2 2 MEXICO CITY 2] 1996-08-25 The leftist [3 3 Popular Revolutionary Army 3] (EPR) in a published report on Sunday said it operated throughout [4 4 Mexico 4] , including the capital , and denied government assertions it was isolated to one state . Commanders " Vicente " and " Oscar " , guarded by a dozen EPR gunmen , said in an interview with [5 5 La Jornada 5] outside [6 6 Mexico City 6] that the armed group was committed to overthrowing the government . " They (government officials) want to present us before public opinion as a local problem , as just being from [7 7 Guerrero 7] and as irrational radicals , " Commander Oscar told [8 8 La Jornada 8] . They said the ERP , whose fighters first appeared wearing military fatigues and brandishing assault rifles in the southwestern state of [9 9 Guerrero 9] on June 28 , had a 23,000-strong membership , but this could not be confirmed independently . [10 10 La Jornada 10] also reported on Sunday that the [11 11 Mexican Army 11] has discovered a 37-page , EPR manual detailing guerrilla tactics and strategies . It quoted the manual as saying : " The objective of the Basic Course on War is to provide for combatants of the EPR basic military knowledge for the armed conflict

against the police and military apparatus of the bourgeoisie . " It was the second time armed " commanders " of the EPR have granted interviews outside [12 12 Guerrero 12] state , an extremely poor and volatile region where leftist protesters often have clashed violently with authorities . Unlike the better known and unrelated [13 13 Zapatista 13] rebels in southeastern [14 14 Chiapas 14] state , the EPR has never taken on the army in direct combat , according to official reports . There have only been a few skirmishes in [1 5 15 Guerrero 15] in which a handful of police , soldiers and civilians have been killed or injured .

1: 3966054 Mexico
2: 18987 Mexico_City
3: 1529105 Popular_Revolutionary_Army
4: 3966054 Mexico
5: 338760 La_Jornada
6: 18987 Mexico_City
7: 195816 Guerrero
8: 338760 La_Jornada
9: 195816 Guerrero
10: 338760 La_Jornada
11: 2948082 Mexican_Army
12: 195816 Guerrero
13: 87066 Zapatista_Army_of_National_Liberation
14: 6787 Chiapas
15: 195816 Guerrero
1: 3966054 Mexico
[pem_pos=1, final_scores=0.6023484,
final_local_scores=0.59720874, pem=-0.3968012,
similarity=24.069479]
2: 18987 Mexico_City
[pem_pos=0, final_scores=0.45430058,
final_local_scores=0.4367252, pem=-0.028392298,
similarity=17.276865]
3: 1529105 Popular_Revolutionary_Army
[pem_pos=0, final_scores=0.48446226,
final_local_scores=0.57816064, pem=0.0, similarity=21.34318]
4: 3966054 Mexico
[pem_pos=1, final_scores=0.54103625,
final_local_scores=0.5302345, pem=-0.18258883,
similarity=20.905914]
5: 338760 La_Jornada
[pem_pos=0, final_scores=0.6178506,
final_local_scores=0.6640776, pem=0.0, similarity=23.907112]
6: 18987 Mexico_City
[pem_pos=0, final_scores=0.49264437,
final_local_scores=0.47860992, pem=-0.028392298,
similarity=18.526789]
7: 195816 Guerrero
[pem_pos=0, final_scores=0.51505214,
final_local_scores=0.48886353, pem=-0.22545685,
similarity=19.904446]

```
8: 338760 La_Jornada
    [pem_pos=0, final_scores=0.55929214,
final_local_scores=0.60011137, pem=0.0, similarity=21.998234]
9: 195816 Guerrero
    [pem_pos=0, final_scores=0.5128668,
final_local_scores=0.48647642, pem=-0.22545685,
similarity=19.833208]
10: 338760 La_Jornada
    [pem_pos=0, final_scores=0.63668805,
final_local_scores=0.6846547, pem=0.0, similarity=24.521175]
11: 2948082 Mexican_Army
    [pem_pos=0, final_scores=0.3251248,
final_local_scores=0.33872345, pem=0.0, similarity=14.197896]
12: 195816 Guerrero
    [pem_pos=0, final_scores=0.53790855,
final_local_scores=0.5138307, pem=-0.22545685,
similarity=20.649515]
13: 87066 Zapatista_Army_of_National_Liberation
    [pem_pos=0, final_scores=0.4127969,
final_local_scores=0.4361238, pem=0.0, similarity=17.104517]
14: 6787 Chiapas
    [pem_pos=0, final_scores=0.4164216,
final_local_scores=0.38032472, pem=0.0, similarity=15.439362]
15: 195816 Guerrero
    [pem_pos=0, final_scores=0.48423463,
final_local_scores=0.45520008, pem=-0.22545685,
similarity=18.89986]
```

Sample: b'380_Ten'

Ten people gunned down in northwest [1 1 Colombia 1] . [2 2 B OGOTA 2] , [3 3 Colombia 3] 1996-08-25 Unidentified gunmen dragged 10 men out of their homes in a rural area of [4 4 Colombia 4] 's northwest province of [5 5 Antioquia 5] and shot them to death , authorities said on Sunday . Police said the killings occurred on Saturday morning in the municipality of Anza but news of the massacre only reached the provincial capital of [6 6 Medellin 6] early on Sunday . Anza is only 20 miles (30 km) west of [7 7 Medellin 7] , but there are no roads linking it directly to the city . Police initially said leftist [8 8 Revolutionary Armed Forces of Colombia 8] ([9 9 FARC 9]) rebels were prime suspects in the killings . But gunmen of the left and right have killed with impunity across [10 10 Antioquia 10] for years , and there were unconfirmed reports that the latest bloodshed was the work of a right-wing paramilitary group .

```
1: 5222 Colombia
2: 211271 Bogotá
3: 5222 Colombia
4: 5222 Colombia
5: 185251 Antioquia_Department
6: 340197 Medellín
7: 340197 Medellín
```

```
8: 26100 FARC
9: 26100 FARC
10: 185251 Antioquia_Department
1: 5222 Colombia
    [pem_pos=0, final_scores=0.6317118,
final_local_scores=0.60587305, pem=-0.22328302,
similarity=23.38442]
2: 211271 Bogotá
    [pem_pos=0, final_scores=0.6513899,
final_local_scores=0.6349978, pem=0.0, similarity=23.039314]
3: 5222 Colombia
    [pem_pos=0, final_scores=0.6049121,
final_local_scores=0.57659847, pem=-0.22328302,
similarity=22.510809]
4: 5222 Colombia
    [pem_pos=0, final_scores=0.6775754,
final_local_scores=0.65597206, pem=-0.22328302,
similarity=24.879475]
5: 185251 Antioquia_Department
    [pem_pos=0, final_scores=0.34278518,
final_local_scores=0.2959719, pem=0.0, similarity=12.922105]
6: 340197 Medellín
    [pem_pos=0, final_scores=0.47169906,
final_local_scores=0.43980026, pem=-0.047747165,
similarity=17.473886]
7: 340197 Medellín
    [pem_pos=0, final_scores=0.44515103,
final_local_scores=0.41080064, pem=-0.047747165,
similarity=16.608479]
8: 26100 FARC
    [pem_pos=0, final_scores=0.5098027,
final_local_scores=0.53807986, pem=0.0, similarity=20.147089]
9: 26100 FARC
    [pem_pos=0, final_scores=0.38819325,
final_local_scores=0.40524, pem=0.0, similarity=16.182882]
10: 185251 Antioquia_Department
    [pem_pos=0, final_scores=0.3644486,
final_local_scores=0.31963596, pem=0.0, similarity=13.628287]
```

Sample: b'381_Port'

Port of Tauranga year profit climbs . [1 1 WELLINGTON 1]
1996-08-26 Year to June 30 . (million [2 2 NZ\$ 2] unless
stated) Net profit 9.050 vs 6.03

```
1: 33804 Wellington
2: 161585 New_Zealand_dollar
1: 33804 Wellington
    [pem_pos=0, final_scores=0.3915642,
final_local_scores=0.45090777, pem=-0.9265605,
similarity=22.58448]
2: 161585 New_Zealand_dollar
    [pem_pos=0, final_scores=0.4299186,
final_local_scores=0.4928041, pem=0.0, similarity=18.79597]
```

Sample: b'382_Two'

Two [1 1 Thai 1] border police wounded by [2 2 Burma 2] gunmen . [3 3 MAE SOT 3] , [4 4 Thailand 4] 1996-08-25 Two [5 5 Thai 5] border policemen were seriously wounded on Sunday when members of a [6 6 Burmese 6] rebel splinter faction ambushed their patrol in northwest [7 7 Thailand 7] , security officers said . The two were wounded in the early hours of Sunday when some 30 members of the [8 8 Democratic Karen Buddhist Army 8] ([9 9 DKBA 9]) ambushed their patrol on the [10 10 Thai 10] side of the border with [11 11 Burma 11] , to the north of the town of [12 12 Mae Sot 12] . The [13 13 Thai 13] army commander in the area , Col Suvit Maenmuang , told reporters the [14 14 DKBA 14] , which is allied with the [15 15 Rangoon 15] military government and based in southeast [16 16 Burma 16] , had recently stepped up cross-border infiltration . Suvit said the motive for their intrusions was not clear but he had ordered reinforcements to beef up security along the porous frontier . The [17 17 DKBA 17] was formed in late 1994 by hundreds of guerrillas who split from the anti-Rangoon [18 18 Karen National Union 18] (KNU) and allied themselves with [19 19 Burmese 19] government army , their former enemies . [20 20 DKBA 20] members have since launched intermittent cross-border attacks on [21 21 Karen 21] refugee camps in [22 22 Thailand 22] , where the majority of inhabitants are KNU supporters , and on [23 23 Thai 23] villages and police posts near the border . [24 24 Bangkok 24] has complained to [25 25 Rangoon 25] about the raids but [26 26 Burmese 26] military authorities say they have no control over the faction . [27 27 Thai 27] army commanders reject the explanation , saying they have evidence the [28 28 Burmese 28] army supplies and directs the renegade ethnic minority splinter faction .

1: 30128 Thailand

2: 19457 Burma

3: 1723280 Mae_Sot_District

4: 30128 Thailand

5: 30128 Thailand

6: 19457 Burma

7: 30128 Thailand

8: 16657986 Democratic_Karen_Buddhist_Army

9: 16657986 Democratic_Karen_Buddhist_Army

10: 30128 Thailand

11: 19457 Burma

12: 1723280 Mae_Sot_District

13: 30128 Thailand

14: 16657986 Democratic_Karen_Buddhist_Army

15: 57404 Yangon

16: 19457 Burma

17: 16657986 Democratic_Karen_Buddhist_Army

18: 504351 Karen_National_Union

19: 19457 Burma

```
20: 16657986 Democratic_Karen_Buddhist_Army
21: 615208 Karen_people
22: 30128 Thailand
23: 30128 Thailand
24: 56615 Bangkok
25: 57404 Yangon
26: 19457 Burma
27: 30128 Thailand
28: 19457 Burma
1: 30128 Thailand
    [pem_pos=2, final_scores=0.63681763,
final_local_scores=0.59957534, pem=-1.7504287,
similarity=31.501339]
2: 19457 Burma
    [pem_pos=0, final_scores=0.7419501,
final_local_scores=0.7081618, pem=-0.20192546,
similarity=26.320776]
3: 1723280 Mae_Sot_District
    [pem_pos=0, final_scores=0.5586122,
final_local_scores=0.6166323, pem=0.0, similarity=22.491251]
4: 30128 Thailand
    [pem_pos=0, final_scores=0.68205065,
final_local_scores=0.64898556, pem=-0.24029273,
similarity=24.763485]
5: 30128 Thailand
    [pem_pos=2, final_scores=0.62404126,
final_local_scores=0.58561903, pem=-1.7504287,
similarity=31.084856]
6: 19457 Burma
    [pem_pos=1, final_scores=0.60300153,
final_local_scores=0.5563816, pem=-1.2286767,
similarity=27.374985]
7: 30128 Thailand
    [pem_pos=0, final_scores=0.6200854,
final_local_scores=0.58129793, pem=-0.24029273,
similarity=22.743551]
8: 16657986 Democratic_Karen_Buddhist_Army
    [pem_pos=0, final_scores=0.54814976,
final_local_scores=0.60237485, pem=0.0, similarity=22.06578]
9: 16657986 Democratic_Karen_Buddhist_Army
    [pem_pos=0, final_scores=0.6017775,
final_local_scores=0.6609551, pem=0.0, similarity=23.81393]
10: 30128 Thailand
    [pem_pos=2, final_scores=0.6309908,
final_local_scores=0.5932104, pem=-1.7504287,
similarity=31.311398]
11: 19457 Burma
    [pem_pos=0, final_scores=0.5856095,
final_local_scores=0.53738344, pem=-0.20192546,
similarity=21.22441]
12: 1723280 Mae_Sot_District
    [pem_pos=0, final_scores=0.572264,
```

```
final_local_scores=0.6315448, pem=0.0, similarity=22.93627]
13: 30128 Thailand
    [pem_pos=2, final_scores=0.65722007,
final_local_scores=0.6218619, pem=-1.7504287,
similarity=32.166412]
14: 16657986 Democratic_Karen_Buddhist_Army
    [pem_pos=0, final_scores=0.6069601,
final_local_scores=0.66661626, pem=0.0, similarity=23.982872]
15: 57404 Yangon
    [pem_pos=0, final_scores=0.42224318,
final_local_scores=0.40042245, pem=0.0, similarity=16.039118]
16: 19457 Burma
    [pem_pos=0, final_scores=0.57531554,
final_local_scores=0.5261388, pem=-0.20192546,
similarity=20.888845]
17: 16657986 Democratic_Karen_Buddhist_Army
    [pem_pos=0, final_scores=0.6173772,
final_local_scores=0.6779953, pem=0.0, similarity=24.322445]
18: 504351 Karen_National_Union
    [pem_pos=0, final_scores=0.5099999,
final_local_scores=0.56364435, pem=0.0, similarity=20.909985]
19: 19457 Burma
    [pem_pos=1, final_scores=0.6913707,
final_local_scores=0.6529115, pem=-1.2286767,
similarity=30.25563]
20: 16657986 Democratic_Karen_Buddhist_Army
    [pem_pos=0, final_scores=0.49644035,
final_local_scores=0.5458902, pem=0.0, similarity=20.380165]
21: 615208 Karen_people
    [pem_pos=0, final_scores=0.4228745,
final_local_scores=0.42239255, pem=-0.7281619,
similarity=20.654606]
22: 30128 Thailand
    [pem_pos=0, final_scores=0.61755913,
final_local_scores=0.5785383, pem=-0.24029273,
similarity=22.661198]
23: 30128 Thailand
    [pem_pos=2, final_scores=0.65847474,
final_local_scores=0.6232324, pem=-1.7504287,
similarity=32.207314]
24: 56615 Bangkok
    [pem_pos=0, final_scores=0.43411642,
final_local_scores=0.4261936, pem=-0.0005963434,
similarity=16.811422]
25: 57404 Yangon
    [pem_pos=0, final_scores=0.38365692,
final_local_scores=0.3582728, pem=0.0, similarity=14.781287]
26: 19457 Burma
    [pem_pos=1, final_scores=0.6094309,
final_local_scores=0.5634047, pem=-1.2286767,
similarity=27.584568]
27: 30128 Thailand
```

```
[pem_pos=2, final_scores=0.6960078,
final_local_scores=0.6642316, pem=-1.7504287,
similarity=33.43081]
28: 19457 Burma
[pem_pos=1, final_scores=0.61293226,
final_local_scores=0.5672294, pem=-1.2286767,
similarity=27.698704]

Sample: b'383_U.S.'
[1 1 U.S. 1] F-14 catches fire while landing in [2 2 Israel 2
] . [3 3 JERUSALEM 3] 1996-08-25 A [4 4 U.S. 4] fighter plane
blew a tyre and caught fire while landing on Sunday at [5 5 Is
rael 5] 's Ben Gurion airport , an airport spokesman said .
A [6 6 U.S. 6] F-14 military plane while landing at Ben
Gurion airport blew a wheel and a fire broke out , " said
spokesman Yehiel Amitai , adding that the two pilots on board
were not injured . " Airport officials declared an emergency
situation at the highest level and the fire brigade put out
the flames while the plane was landing , " he said .
1: 3434750 United_States
2: 9282173 Israel
3: 16043 Jerusalem
4: 3434750 United_States
5: 9282173 Israel
6: 3434750 United_States
1: 3434750 United_States
[pem_pos=0, final_scores=0.4907729,
final_local_scores=0.44498914, pem=-0.14075686,
similarity=18.134533]
2: 9282173 Israel
[pem_pos=0, final_scores=0.5332536,
final_local_scores=0.49493867, pem=-0.21279564,
similarity=20.016886]
3: 16043 Jerusalem
[pem_pos=0, final_scores=0.48866886,
final_local_scores=0.5064192, pem=-0.13239029,
similarity=19.92223]
4: 3434750 United_States
[pem_pos=0, final_scores=0.5184956,
final_local_scores=0.475272, pem=-0.14075686,
similarity=19.038235]
5: 9282173 Israel
[pem_pos=0, final_scores=0.59718233,
final_local_scores=0.5647711, pem=-0.21279564,
similarity=22.100826]
6: 3434750 United_States
[pem_pos=0, final_scores=0.5108333,
final_local_scores=0.46690208, pem=-0.14075686,
similarity=18.78846]
```

Sample: b'384_Egypt'
[1 1 Egypt 1] hopes jars will reveal secrets of mummies . [2

2 CAIRO 2] 1996-08-25 Archaeologists in [3 3 Egypt 3] have found pots used by ancient [4 4 Egyptians 4] in burial rites that they say may reveal the secrets of mummification . Mohammed Saleh , director of the [5 5 Egyptian 5] Museum , told [6 6 Reuters 6] television a [7 7 U.S. 7] team found the pots , some of which contain human intestines , in a tomb built into the rocks while digging in Dahshour , a village 40 km (25 miles) south of [8 8 Cairo 8] . Dahshour is the site of [9 9 Egypt 9] 's second largest pyramid , built for the pharaoh Seneferu more than 4,500 years ago . Saleh said the team -- from [10 10 New York 10] 's [11 11 Metropolitan Museum 11] -- found four [12 12 Canopic 12] jars and two unguent jars in the tomb , which belongs to an unidentified person who lived during the 12th Dynasty (1991-1786 BC) in the [13 13 Middle Kingdom 13] . " This finding is important because one of the jars still contains substances and materials used in the conservation of mummies and the conservation of the intestines and all the things which were in the cavity of a person we have not identified yet , " Saleh said . " We hope that the analysis of such substances and liquids will reveal some secrets of the mummification process and materials used in this process , " he added .

1: 8087628 Egypt

2: 6293 Cairo

3: 8087628 Egypt

4: 31912046 Egyptians

5: 8087628 Egypt

6: 18998750 Reuters

7: 3434750 United_States

8: 6293 Cairo

9: 8087628 Egypt

10: 645042 New_York_City

11: 37535 Metropolitan_Museum_of_Art

12: 235539 Canopic_jar

13: 292602 Middle_Kingdom_of_Egypt

1: 8087628 Egypt

[pem_pos=0, final_scores=0.50811666, final_local_scores=0.4521296, pem=-0.27480945, similarity=19.076618]

2: 6293 Cairo

[pem_pos=0, final_scores=0.63664156, final_local_scores=0.60245955, pem=-0.15211283, similarity=22.895521]

3: 8087628 Egypt

[pem_pos=0, final_scores=0.4966728, final_local_scores=0.4396289, pem=-0.27480945, similarity=18.703573]

4: 31912046 Egyptians

[pem_pos=0, final_scores=0.18243264, final_local_scores=0.1486367, pem=-0.407254, similarity=10.740039]

5: 8087628 Egypt

```
[pem_pos=0, final_scores=0.4163506,
final_local_scores=0.35188913, pem=-1.052751,
similarity=20.31581]
6: 18998750 Reuters
    [pem_pos=0, final_scores=0.41035402,
final_local_scores=0.44919866, pem=0.0, similarity=17.494698]
7: 3434750 United_States
    [pem_pos=0, final_scores=0.47932726,
final_local_scores=0.49771732, pem=-0.14075686,
similarity=19.70805]
8: 6293 Cairo
    [pem_pos=0, final_scores=0.56505805,
final_local_scores=0.5242654, pem=-0.15211283,
similarity=20.562052]
9: 8087628 Egypt
    [pem_pos=0, final_scores=0.5033956,
final_local_scores=0.4469726, pem=-0.27480945,
similarity=18.922724]
10: 645042 New_York_City
    [pem_pos=1, final_scores=0.4439562,
final_local_scores=0.47032434, pem=-1.7906687,
similarity=27.863064]
11: 37535 Metropolitan_Museum_of_Art
    [pem_pos=0, final_scores=0.4103518,
final_local_scores=0.4704334, pem=0.0, similarity=18.128384]
12: 235539 Canopic_jar
    [pem_pos=1, final_scores=0.26691335,
final_local_scores=0.35821304, pem=-0.07109595,
similarity=15.166135]
13: 292602 Middle_Kingdom_of_Egypt
    [pem_pos=0, final_scores=0.3519771,
final_local_scores=0.3882512, pem=0.0, similarity=15.675902]
```

Sample: b'385_Saudi'

```
[1 1 Saudi Arabia 1] executes [2 2 Pakistani 2] man . [3 3 DU
BAI 3] 1996-08-25 [4 4 Saudi Arabia 4] executed on Sunday a [5
5 Pakistani 5] man accused of belonging to an armed gang of
robbers , [6 6 Saudi 6] television reported . It quoted an
Interior Ministry statement as saying Shabir Ahmad Muhammad
Jalil was executed in [7 7 Mecca 7] . He was the 26th person
executed this year in the kingdom . [8 8 Saudi Arabia 8]
beheads convicted drug smugglers , rapists , murderers and
other criminals .
```

```
1: 349303 Saudi_Arabia
2: 23235 Pakistan
3: 211583 Dubai
4: 349303 Saudi_Arabia
5: 23235 Pakistan
6: 349303 Saudi_Arabia
7: 21021 Mecca
8: 349303 Saudi_Arabia
1: 349303 Saudi_Arabia
```

```
[pem_pos=0, final_scores=0.5346855,
final_local_scores=0.46796304, pem=-0.25454172,
similarity=19.4389]
2: 23235 Pakistan
[pem_pos=0, final_scores=0.58950865,
final_local_scores=0.56841856, pem=-0.49826378,
similarity=23.76209]
3: 211583 Dubai
[pem_pos=0, final_scores=0.535348,
final_local_scores=0.54318583, pem=-0.08179128,
similarity=20.744255]
4: 349303 Saudi_Arabia
[pem_pos=0, final_scores=0.5644718,
final_local_scores=0.5005, pem=-0.25454172,
similarity=20.409868]
5: 23235 Pakistan
[pem_pos=0, final_scores=0.54307777,
final_local_scores=0.5176999, pem=-0.49826378,
similarity=22.248547]
6: 349303 Saudi_Arabia
[pem_pos=0, final_scores=0.65958065,
final_local_scores=0.604392, pem=0.0, similarity=22.125975]
7: 21021 Mecca
[pem_pos=0, final_scores=0.35269624,
final_local_scores=0.35471055, pem=0.0, similarity=14.674983]
8: 349303 Saudi_Arabia
[pem_pos=0, final_scores=0.50426996,
final_local_scores=0.4347387, pem=-0.25454172,
similarity=18.447422]
```

Sample: b'386_PRESS'

PRESS DIGEST - [1 1 Jordan 1] - Aug 25 . [2 2 AMMAN 2]
1996-08-25 These are some of the leading stories in the [3 3 J
ordanian 3] press on Sunday . [4 4 Reuters 4] has not
verified these stories and does not vouch for their
accuracy . [5 5 JORDAN TIMES 5] - King : [6 6 Jordan 6] is
entering a new era . No going back on democracy ; attempts to
tamper with security and stability will not be tolerated .
Information Minister Marwan Muasher says there is evidence
that " some official parties in [7 7 Iraq 7] " were behind
the disturbances in the south . - King to visit [8 8 Bahrain 8
] soon . - Government asks senior [9 9 Iraqi 9] " diplomat "
to leave , reviews status of others . - [10 10 Japanese 10]
foreign minister arrives for talks on peace process ,
bilateral ties . [11 11 AL RAI 11] - Prime Minister Abdul-
Karim al-Kabariti says government committed to lifting ceiling
of democracy . - [12 12 Saudi 12] Prince Sultan telephones
prime minister . - [13 13 Jordan 13] releases 32 from
southern town of [14 14 Karak 14] . - [15 15 Jordan 15]
expresses anger at conduct of some [16 16 Iraqi 16] diplomats
in [17 17 Amman 17] which are incompatible with diplomatic
traditions . AD DUSTOUR - Kabariti and parliament speaker

meet to discuss ways to reactivate parliament 's legislative role . AL ASWAQ - State security court starts investigating suspects in unrest .

1: 7515964 Jordan
2: 45007 Amman
3: 7515964 Jordan
4: 18998750 Reuters
5: 3150792 The_Jordan_Times
6: 7515964 Jordan
7: 7515928 Iraq
8: 18933277 Bahrain
9: 7515928 Iraq
10: 15573 Japan
11: 18968435 Al_Rai
12: 349303 Saudi_Arabia
13: 7515964 Jordan
14: 307576 Al_Karak
15: 7515964 Jordan
16: 7515928 Iraq
17: 45007 Amman
1: 7515964 Jordan
[pem_pos=0, final_scores=0.6384931,
final_local_scores=0.5853369, pem=-0.53647494,
similarity=24.474768]
2: 45007 Amman
[pem_pos=0, final_scores=0.54193574,
final_local_scores=0.5090851, pem=0.0, similarity=19.281828]
3: 7515964 Jordan
[pem_pos=0, final_scores=0.5404704,
final_local_scores=0.47826207, pem=-0.18608934,
similarity=19.37399]
4: 18998750 Reuters
[pem_pos=0, final_scores=0.32851762,
final_local_scores=0.34879246, pem=0.0, similarity=14.498376]
5: 3150792 The_Jordan_Times
[pem_pos=0, final_scores=0.47810954,
final_local_scores=0.5731529, pem=0.0, similarity=21.193739]
6: 7515964 Jordan
[pem_pos=0, final_scores=0.50279886,
final_local_scores=0.43711162, pem=-0.53647494,
similarity=20.05143]
7: 7515928 Iraq
[pem_pos=0, final_scores=0.58244544,
final_local_scores=0.53953606, pem=-0.25379938,
similarity=21.570745]
8: 18933277 Bahrain
[pem_pos=0, final_scores=0.40033484,
final_local_scores=0.35154092, pem=-0.42385748,
similarity=16.885397]
9: 7515928 Iraq
[pem_pos=0, final_scores=0.5700523,
final_local_scores=0.5259984, pem=-0.48469657,

```
similarity=22.422411]
10: 15573 Japan
    [pem_pos=0, final_scores=0.7879043,
final_local_scores=0.8607326, pem=-0.99076414,
similarity=35.16362]
11: 18968435 Al_Rai
    [pem_pos=0, final_scores=0.53078693,
final_local_scores=0.66747713, pem=0.0, similarity=24.008562]
12: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.6202348,
final_local_scores=0.5798488, pem=0.0, similarity=21.393559]
13: 7515964 Jordan
    [pem_pos=0, final_scores=0.56080616,
final_local_scores=0.50047576, pem=-0.53647494,
similarity=21.942343]
14: 307576 Al_Karak
    [pem_pos=0, final_scores=0.5082241,
final_local_scores=0.54795426, pem=-0.6855135,
similarity=24.169691]
15: 7515964 Jordan
    [pem_pos=0, final_scores=0.67721313,
final_local_scores=0.62763274, pem=-0.53647494,
similarity=25.73696]
16: 7515928 Iraq
    [pem_pos=0, final_scores=0.5599994,
final_local_scores=0.5150171, pem=-0.48469657,
similarity=22.094706]
17: 45007 Amman
    [pem_pos=0, final_scores=0.42301017,
final_local_scores=0.37917703, pem=0.0, similarity=15.405112]
```

Sample: b'387_Netanyahu'

[1 1 Netanyahu 1] , [2 2 Weizman 2] consult on [3 3 Arafat 3] invitation . [4 4 JERUSALEM 4] 1996-08-25 [5 5 Israeli 5] President [6 6 Ezer Weizman 6] , weighing a possible meeting with [7 7 Yasser Arafat 7] , consulted on Sunday with Prime Minister [8 8 Benjamin Netanyahu 8] , a spokesman said . [9 9 Weizman 9] and [10 10 Netanyahu 10] met at the president 's official [11 11 Jerusalem 11] residence and planned to speak to the media at the end of their talks , the prime minister 's spokesman said . Earlier , the director of the president 's office denied a report in [12 12 Israel 12] 's [13 13 Yedioth Ahronoth 13] newspaper that [14 14 Weizman 14] had already invited [15 15 Arafat 15] to his private home for talks in the coming week on the future of the Israel-PLO peace process . But the official , Aryeh Shumer , said it was only fitting that [16 16 Weizman 16] and [17 17 Arafat 17] should talk after the [18 18 Palestinian 18] leader sent the [19 19 Israeli 19] president a letter which [20 20 Yedioth Ahronoth 20] reported contained an emotional appeal to save the peace proces . The newspaper said [21 21 Netanyahu 21] , who is cool to meeting [22 22 Arafat 22] himself , opposed talks

between [23 23 Weizman 23] and the [24 24 Palestinian 24] president . After [25 25 Moslem 25] suicide bombers killed 59 people in [26 26 Israel 26] in February and March , [27 27 Weizman 27] called for peace efforts with the PLO to be suspended . Shumer said his current position was that the peace process must continue .

1: 56472 Benjamin_Netanyahu
2: 673848 Ezer_Weizman
3: 34289 Yasser_Arafat
4: 16043 Jerusalem
5: 9282173 Israel
6: 673848 Ezer_Weizman
7: 34289 Yasser_Arafat
8: 56472 Benjamin_Netanyahu
9: 673848 Ezer_Weizman
10: 56472 Benjamin_Netanyahu
11: 16043 Jerusalem
12: 9282173 Israel
13: 450450 Yedioth_Ahronoth
14: 673848 Ezer_Weizman
15: 34289 Yasser_Arafat
16: 673848 Ezer_Weizman
17: 34289 Yasser_Arafat
18: 241405 State_of_Palestine
19: 9282173 Israel
20: 450450 Yedioth_Ahronoth
21: 56472 Benjamin_Netanyahu
22: 34289 Yasser_Arafat
23: 673848 Ezer_Weizman
24: 241405 State_of_Palestine
25: 6037917 Islam
26: 9282173 Israel
27: 673848 Ezer_Weizman
1: 56472 Benjamin_Netanyahu
 [pem_pos=0, final_scores=0.5195624,
final_local_scores=0.4736662, pem=0.0, similarity=18.224857]
2: 673848 Ezer_Weizman
 [pem_pos=0, final_scores=0.39527518,
final_local_scores=0.37022156, pem=0.0, similarity=15.137863]
3: 34289 Yasser_Arafat
 [pem_pos=0, final_scores=0.5990365,
final_local_scores=0.56024104, pem=0.0, similarity=20.808424]
4: 16043 Jerusalem
 [pem_pos=0, final_scores=0.5131783,
final_local_scores=0.50943595, pem=-0.13239029,
similarity=20.012257]
5: 9282173 Israel
 [pem_pos=0, final_scores=0.5313409,
final_local_scores=0.5109454, pem=-0.34790626,
similarity=21.22931]
6: 673848 Ezer_Weizman
 [pem_pos=0, final_scores=0.5907907,

```
final_local_scores=0.58379257, pem=0.0, similarity=21.511248]
7: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.5784604,
final_local_scores=0.5355258, pem=0.0, similarity=20.070871]
8: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.73232627,
final_local_scores=0.70382434, pem=0.0, similarity=25.093235]
9: 673848 Ezer_Weizman
    [pem_pos=0, final_scores=0.44669628,
final_local_scores=0.4263913, pem=0.0, similarity=16.81408]
10: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.53888685,
final_local_scores=0.49477512, pem=0.0, similarity=18.85479]
11: 16043 Jerusalem
    [pem_pos=0, final_scores=0.41005754,
final_local_scores=0.39679217, pem=-0.13239029,
similarity=16.650742]
12: 9282173 Israel
    [pem_pos=0, final_scores=0.70738894,
final_local_scores=0.7032512, pem=-0.21279564,
similarity=26.233345]
13: 450450 Yedioth_Ahronoth
    [pem_pos=0, final_scores=0.5379601,
final_local_scores=0.5583814, pem=0.0, similarity=20.752928]
14: 673848 Ezer_Weizman
    [pem_pos=0, final_scores=0.50746614,
final_local_scores=0.49277318, pem=0.0, similarity=18.795048]
15: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.6221215,
final_local_scores=0.5854579, pem=0.0, similarity=21.560946]
16: 673848 Ezer_Weizman
    [pem_pos=0, final_scores=0.5442949,
final_local_scores=0.533003, pem=0.0, similarity=19.995584]
17: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.68137366,
final_local_scores=0.6501819, pem=0.0, similarity=23.492437]
18: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.5574456,
final_local_scores=0.5730059, pem=-2.8508017,
similarity=36.692455]
19: 9282173 Israel
    [pem_pos=0, final_scores=0.55435044,
final_local_scores=0.5360798, pem=-0.34790626,
similarity=21.979372]
20: 450450 Yedioth_Ahronoth
    [pem_pos=0, final_scores=0.5620938,
final_local_scores=0.5847438, pem=0.0, similarity=21.539635]
21: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.567816,
final_local_scores=0.5263759, pem=0.0, similarity=19.79782]
22: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.62637097,
```

```
final_local_scores=0.5900998, pem=0.0, similarity=21.699469]
23: 673848 Ezer_Weizman
    [pem_pos=0, final_scores=0.49569857,
final_local_scores=0.47991884, pem=0.0, similarity=18.411448]
24: 241405 State_of_Palestine
    [pem_pos=10, final_scores=0.64496654,
final_local_scores=0.66860926, pem=-2.8508017,
similarity=39.545452]
25: 6037917 Islam
    [pem_pos=1, final_scores=0.35727203,
final_local_scores=0.42744732, pem=-1.2150909,
similarity=23.453445]
26: 9282173 Israel
    [pem_pos=0, final_scores=0.59937507,
final_local_scores=0.58526236, pem=-0.21279564,
similarity=22.712326]
27: 673848 Ezer_Weizman
    [pem_pos=0, final_scores=0.51371974,
final_local_scores=0.49960423, pem=0.0, similarity=18.9989]
```

Sample: b'388_PRESS'

PRESS DIGEST - [1 1 Israel 1] - Aug 25 . [2 2 JERUSALEM 2]
1996-08-25 These are some of the leading stories in [3 3 Isra
eli 3] newspapers on Sunday . [4 4 Reuters 4] has not
verified these stories and does not vouch for their
accuracy . [5 5 HAARETZ 5] - [6 6 Palestinian 6] President [7
7 Arafat 7] opens civilian struggle against [8 8 Israel 8] ,
calls on [9 9 Palestinians 9] to build in self-rule areas . -
Seven ministers and governor of [10 10 Bank of Israel 10]
will visit the [11 11 United States 11] at the end of
September and in October . - [12 12 Israel 12] bans plane
donated by the [13 13 Netherlands 13] to [14 14 Arafat 14] to
land at [15 15 Gaza 15] airport . - Former prime minister [16
16 Peres 16] to [17 17 Morocco 17] today . [18 18 YEDIOTH AHR
ONOTH 18] - [19 19 Israeli 19] President [20 20 Weizman 20]
invited [21 21 Palestinian 21] President [22 22 Arafat 22] to
meet him at his private residence . - [23 23 Netanyahu 23]
opposes transit camps for foreign workers facing expulsion .
- Foreign Minister Levy to visit [24 24 Egypt 24] soon . [25 2
5 MAARIV 25] - [26 26 Palestinian Authority 26] has taken
over education in [27 27 East Jerusalem 27] . - [28 28 Syrian
28] armoured columns on the move in [29 29 Lebanon 29] . - [30
30 Shimon Peres 30] to [31 31 Morocco 31] , will stay at king
's private residence . [32 32 JERUSALEM POST 32] - [33 33 Pal
estinian 33] Minister [34 34 Erekat 34] says Israel-PLO talks
will begin by September 2 . - Prime minister names former
general Avraham Tamir to staff after failing to establish
national security council . - Cabinet puts off decision on
foreign workers . - Internal Security Minister [35 35 Kahalani
35] warns cabinet of increase in organised crime .

1: 9282173 Israel
2: 16043 Jerusalem

3: 9282173 Israel
4: 18998750 Reuters
5: 64878 Haaretz
6: 241405 State_of_Palestine
7: 34289 Yasser_Arafat
8: 9282173 Israel
9: 1114732 Palestine
10: 736051 Bank_of_Israel
11: 3434750 United_States
12: 9282173 Israel
13: 21148 Netherlands
14: 34289 Yasser_Arafat
15: 55387 Gaza
16: 48569 Shimon_Peres
17: 19291 Morocco
18: 450450 Yedioth_Ahronoth
19: 9282173 Israel
20: 673848 Ezer_Weizman
21: 1114732 Palestine
22: 34289 Yasser_Arafat
23: 56472 Benjamin_Netanyahu
24: 8087628 Egypt
25: 374605 Maariv_(newspaper)
26: 24093 Palestinian_National_Authority
27: 343508 East_Jerusalem
28: 7515849 Syria
29: 17771 Lebanon
30: 48569 Shimon_Peres
31: 19291 Morocco
32: 184764 The_Jerusalem_Post
33: 241405 State_of_Palestine
34: 556603 Saeb_Erekat
35: 5952590 Avigdor_Khalani
1: 9282173 Israel
 [pem_pos=0, final_scores=0.5475571,
final_local_scores=0.50101596, pem=-0.21279564,
similarity=20.198244]
2: 16043 Jerusalem
 [pem_pos=0, final_scores=0.53032285,
final_local_scores=0.50922686, pem=-0.13239029,
similarity=20.006018]
3: 9282173 Israel
 [pem_pos=0, final_scores=0.5566385,
final_local_scores=0.51093596, pem=-0.34790626,
similarity=21.22903]
4: 18998750 Reuters
 [pem_pos=0, final_scores=0.35298747,
final_local_scores=0.3585363, pem=0.0, similarity=14.789151]
5: 64878 Haaretz
 [pem_pos=0, final_scores=0.57301015,
final_local_scores=0.5741337, pem=0.0, similarity=21.223007]
6: 241405 State_of_Palestine

```
[pem_pos=9, final_scores=0.59984285,
final_local_scores=0.6043881, pem=-2.8508017,
similarity=37.628963]
7: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.60494417,
final_local_scores=0.5781587, pem=0.0, similarity=21.34312]
8: 9282173 Israel
    [pem_pos=0, final_scores=0.5703066,
final_local_scores=0.5258663, pem=-0.21279564,
similarity=20.939829]
9: 1114732 Palestine
    [pem_pos=1, final_scores=0.3545395,
final_local_scores=0.30961436, pem=-1.6411982,
similarity=22.254314]
10: 736051 Bank_of_Israel
    [pem_pos=0, final_scores=0.49702746,
final_local_scores=0.5646837, pem=0.0, similarity=20.941]
11: 3434750 United_States
    [pem_pos=0, final_scores=0.39652032,
final_local_scores=0.42097336, pem=-0.073457584,
similarity=17.051872]
12: 9282173 Israel
    [pem_pos=0, final_scores=0.6194256,
final_local_scores=0.57952136, pem=-0.21279564,
similarity=22.541]
13: 21148 Netherlands
    [pem_pos=0, final_scores=0.40956366,
final_local_scores=0.46188772, pem=-0.24940795,
similarity=19.229683]
14: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.53407073,
final_local_scores=0.5007402, pem=0.0, similarity=19.032799]
15: 55387 Gaza
    [pem_pos=0, final_scores=0.29403454,
final_local_scores=0.24561787, pem=-0.30319342,
similarity=13.0682535]
16: 48569 Shimon_Peres
    [pem_pos=0, final_scores=0.658184,
final_local_scores=0.6625935, pem=0.0, similarity=23.862823]
17: 19291 Morocco
    [pem_pos=0, final_scores=0.6237692,
final_local_scores=0.65468025, pem=-0.29598162,
similarity=25.23627]
18: 450450 Yedioth_Ahronoth
    [pem_pos=0, final_scores=0.5578844,
final_local_scores=0.5753716, pem=0.0, similarity=21.25995]
19: 9282173 Israel
    [pem_pos=0, final_scores=0.5320177,
final_local_scores=0.4840415, pem=-0.34790626,
similarity=20.426447]
20: 673848 Ezer_Weizman
    [pem_pos=0, final_scores=0.41942507,
```

```
final_local_scores=0.4447251, pem=0.0, similarity=17.361197]
21: 1114732 Palestine
    [pem_pos=1, final_scores=0.26873535,
final_local_scores=0.21588635, pem=-1.9100745,
similarity=20.919474]
22: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.56559145,
final_local_scores=0.53517175, pem=0.0, similarity=20.060305]
23: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.49398196,
final_local_scores=0.46950722, pem=-0.021035189,
similarity=18.215137]
24: 8087628 Egypt
    [pem_pos=0, final_scores=0.52279824,
final_local_scores=0.51380175, pem=-0.27480945,
similarity=20.91704]
25: 374605 Maariv_(newspaper)
    [pem_pos=1, final_scores=0.4068777,
final_local_scores=0.46213126, pem=-0.62108535,
similarity=21.25819]
26: 24093 Palestinian_National_Authority
    [pem_pos=0, final_scores=0.42131442,
final_local_scores=0.38321382, pem=0.0, similarity=15.525578]
27: 343508 East_Jerusalem
    [pem_pos=0, final_scores=0.39654237,
final_local_scores=0.3786316, pem=0.0, similarity=15.388834]
28: 7515849 Syria
    [pem_pos=0, final_scores=0.5402332,
final_local_scores=0.5405917, pem=-0.60890055,
similarity=23.533344]
29: 17771 Lebanon
    [pem_pos=0, final_scores=0.5056885,
final_local_scores=0.4854532, pem=-0.4097839,
similarity=20.805075]
30: 48569 Shimon_Peres
    [pem_pos=0, final_scores=0.64501506,
final_local_scores=0.6441866, pem=0.0, similarity=23.313526]
31: 19291 Morocco
    [pem_pos=0, final_scores=0.54289675,
final_local_scores=0.5663394, pem=-0.29598162,
similarity=22.600002]
32: 184764 The_Jerusalem_Post
    [pem_pos=0, final_scores=0.43306845,
final_local_scores=0.42771387, pem=0.0, similarity=16.853548]
33: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.66959625,
final_local_scores=0.680583, pem=-2.8508017,
similarity=39.90277]
34: 556603 Saeb_Erekat
    [pem_pos=0, final_scores=0.45021123,
final_local_scores=0.4709108, pem=0.0, similarity=18.14263]
35: 5952590 Avigdor_Kahalani
```

```
[pem_pos=0, final_scores=0.40456152,
final_local_scores=0.48806798, pem=0.0, similarity=18.654634]
```

Sample: b'389_Voting'

Voting begins in second round of [1 1 Lebanese 1] election . [2 2 TRIPOLI 2] , [3 3 Lebanon 3] 1996-08-25 Voting began on Sunday in north [4 4 Lebanon 4] in the second round of parliamentary elections with 580,000 voters eligible to choose 28 members of the 128-member parliament . A thin trickle of voters began casting their ballots in this northern port city for the five rival lists of candidates as polling stations opened at 7 a.m. (0400 gmt) .

1: 17771 Lebanon

2: 268796 Tripoli,_Lebanon

3: 17771 Lebanon

4: 17771 Lebanon

1: 17771 Lebanon

```
[pem_pos=0, final_scores=0.65531933,
final_local_scores=0.60561776, pem=-0.55808944,
similarity=25.197533]
```

2: 268796 Tripoli,_Lebanon

```
[pem_pos=1, final_scores=0.23901702,
final_local_scores=0.22196281, pem=-1.8286858,
similarity=20.658203]
```

3: 17771 Lebanon

```
[pem_pos=0, final_scores=0.58185726,
final_local_scores=0.5253716, pem=-0.4097839,
similarity=21.996317]
```

4: 17771 Lebanon

```
[pem_pos=0, final_scores=0.510119,
final_local_scores=0.44700843, pem=-0.4097839,
similarity=19.657806]
```

Sample: b'390_Israeli'

[1 1 Israeli 1] president invites [2 2 Arafat 2] to home - paper . [3 3 JERUSALEM 3] 1996-08-25 [4 4 Israeli 4] President [5 5 Ezer Weizman 5] has invited [6 6 Yasser Arafat 6] to meet him at his private home , [7 7 Israel 7] 's biggest newspaper said on Sunday . The [8 8 Yedioth Ahronoth 8] daily reported that Prime Minister [9 9 Benjamin Netanyahu 9] , who has said he has no desire to hold talks with the [10 10 Palestinian 10] president , opposes the meeting due to be held this coming week . The newspaper said [11 11 Weizman 11] scheduled the meeting at his private residence in the central [12 12 Israeli 12] village of [13 13 Caesarea 13] after [14 14 Arafat 14] sent him an emotional appeal " to save the peace process " . [15 15 Netanyahu 15] met [16 16 Weizman 16] last Tuesday and voiced his opposition , [17 17 Yedioth 17] said . " I am prepared to postpone the meeting under one condition -- that you give me a commitment right now to meet [18 18 Arafat 18] yourself within 10 days , " the paper quoted [19 19 Weizman 19] as telling [20 20 Netanyahu 20] . It said [21 21 Ne

tanyahu 21] had yet to give [22 22 Weizman 22] an answer .
The office of [23 23 Israeli 23] president is largely
ceremonial . But [24 24 Weizman 24] , a former defence
minister and an architect of [25 25 Israel 25] 's peace
treaty with [26 26 Egypt 26] , has spoken out frequently on
the peace process with the [27 27 Palestinians 27] -- at
times urging the former [28 28 Labour 28] government to slow
it down .

1: 9282173 Israel
2: 34289 Yasser_Arafat
3: 16043 Jerusalem
4: 9282173 Israel
5: 673848 Ezer_Weizman
6: 34289 Yasser_Arafat
7: 9282173 Israel
8: 450450 Yedioth_Ahronoth
9: 56472 Benjamin_Netanyahu
10: 241405 State_of_Palestine
11: 673848 Ezer_Weizman
12: 9282173 Israel
13: 3530849 Caesarea
14: 34289 Yasser_Arafat
15: 56472 Benjamin_Netanyahu
16: 673848 Ezer_Weizman
17: 450450 Yedioth_Ahronoth
18: 34289 Yasser_Arafat
19: 673848 Ezer_Weizman
20: 56472 Benjamin_Netanyahu
21: 56472 Benjamin_Netanyahu
22: 673848 Ezer_Weizman
23: 9282173 Israel
24: 673848 Ezer_Weizman
25: 9282173 Israel
26: 8087628 Egypt
27: 241405 State_of_Palestine
28: 40312187 Israeli_Labor_Party
1: 9282173 Israel
 [pem_pos=0, final_scores=0.5767345,
final_local_scores=0.55017775, pem=-0.34790626,
similarity=22.400084]
2: 34289 Yasser_Arafat
 [pem_pos=0, final_scores=0.6329556,
final_local_scores=0.6037387, pem=0.0, similarity=22.106482]
3: 16043 Jerusalem
 [pem_pos=0, final_scores=0.53106064,
final_local_scores=0.5257477, pem=-0.13239029,
similarity=20.499033]
4: 9282173 Israel
 [pem_pos=0, final_scores=0.52088857,
final_local_scores=0.48917454, pem=-0.34790626,
similarity=20.579626]
5: 673848 Ezer_Weizman

```
[pem_pos=0, final_scores=0.5793697,
final_local_scores=0.57547456, pem=0.0, similarity=21.263021]
6: 34289 Yasser_Arafat
[pem_pos=0, final_scores=0.5410898,
final_local_scores=0.5007356, pem=0.0, similarity=19.032661]
7: 9282173 Israel
[pem_pos=0, final_scores=0.585791,
final_local_scores=0.5600706, pem=-0.21279564,
similarity=21.96055]
8: 450450 Yedioth_Ahronoth
[pem_pos=0, final_scores=0.47584075,
final_local_scores=0.48733538, pem=0.0, similarity=18.632772]
9: 56472 Benjamin_Netanyahu
[pem_pos=0, final_scores=0.7022159,
final_local_scores=0.67142665, pem=0.0, similarity=24.126423]
10: 241405 State_of_Palestine
[pem_pos=9, final_scores=0.6538586,
final_local_scores=0.67572415, pem=-2.8508017,
similarity=39.757774]
11: 673848 Ezer_Weizman
[pem_pos=0, final_scores=0.50034493,
final_local_scores=0.48915195, pem=0.0, similarity=18.686983]
12: 9282173 Israel
[pem_pos=0, final_scores=0.53866386,
final_local_scores=0.50859135, pem=-0.34790626,
similarity=21.159061]
13: 3530849 Caesarea
[pem_pos=0, final_scores=0.24727468,
final_local_scores=0.2795336, pem=-0.38856283,
similarity=14.544618]
14: 34289 Yasser_Arafat
[pem_pos=0, final_scores=0.5547905,
final_local_scores=0.51835525, pem=0.0, similarity=19.558468]
15: 56472 Benjamin_Netanyahu
[pem_pos=0, final_scores=0.5595817,
final_local_scores=0.5179194, pem=0.0, similarity=19.545462]
16: 673848 Ezer_Weizman
[pem_pos=0, final_scores=0.4666497,
final_local_scores=0.452345, pem=0.0, similarity=17.58859]
17: 450450 Yedioth_Ahronoth
[pem_pos=0, final_scores=0.46667546,
final_local_scores=0.47732365, pem=0.0, similarity=18.334003]
18: 34289 Yasser_Arafat
[pem_pos=0, final_scores=0.6226774,
final_local_scores=0.59251136, pem=0.0, similarity=21.771433]
19: 673848 Ezer_Weizman
[pem_pos=0, final_scores=0.50472355,
final_local_scores=0.49393493, pem=0.0, similarity=18.829716]
20: 56472 Benjamin_Netanyahu
[pem_pos=0, final_scores=0.53532296,
final_local_scores=0.49142045, pem=0.0, similarity=18.754679]
21: 56472 Benjamin_Netanyahu
```

```
[pem_pos=0, final_scores=0.52131754,
final_local_scores=0.47612166, pem=0.0, similarity=18.298134]
22: 673848 Ezer_Weizman
    [pem_pos=0, final_scores=0.45531267,
final_local_scores=0.43996102, pem=0.0, similarity=17.219027]
23: 9282173 Israel
    [pem_pos=0, final_scores=0.5867462,
final_local_scores=0.561114, pem=-0.34790626,
similarity=22.726442]
24: 673848 Ezer_Weizman
    [pem_pos=0, final_scores=0.49094993,
final_local_scores=0.47888935, pem=0.0, similarity=18.380726]
25: 9282173 Israel
    [pem_pos=0, final_scores=0.6205625,
final_local_scores=0.59805316, pem=-0.21279564,
similarity=23.094028]
26: 8087628 Egypt
    [pem_pos=0, final_scores=0.53736895,
final_local_scores=0.56235725, pem=-0.27480945,
similarity=22.366032]
27: 241405 State_of_Palestine
    [pem_pos=5, final_scores=0.5956997,
final_local_scores=0.6121944, pem=-1.7909594,
similarity=32.09833]
28: 40312187 Israeli_Labor_Party
    [pem_pos=7, final_scores=0.46316993,
final_local_scores=0.45911652, pem=-3.1950269,
similarity=35.165718]
```

Sample: b'391_Corporate'

Corporate [1 1 America 1] taking new view on compensation . Anne Murray [2 2 NEW YORK 2] 1996-08-23 Corporate [3 3 America 3] is planning major changes in employee compensation in the next few years , according to a recent study . What it comes down to is this : If you 're highly skilled , you 'll benefit nicely . But if you 're not and cannot contribute to your employer 's goals , you 'll be paid less . The survey , conducted in late 1995 and the early part of this year by management consulting firm [4 4 Towers Perrin 4] , showed that the focus will be on an employee 's overall value to the company 's bottom line -- rather than how well an employee performs a specific task . Presently , for example , if an accountant 's job involves doing five specific tasks , he or she can expect a certain salary , said Sandra O'Neal , a [5 5 Towers Perrin 5] principal . In the future , the accountant will be evaluated solely on " knowledge , skill and abilities , " she said . In addition to using accounting skills , the accountant will also have to be creative , work well in a team , be sensitive to customer needs and set productivity goals . " The good news is , if you 're highly skilled and have many abilities , you 'll be paid more , " said O'Neal . " The bad news is , if you 're not skilled and

can't contribute to a team, to customer service and the organisation's goals, you'll be paid less." Of the 750 mid-to-large size corporations surveyed, 81 percent had undergone a major restructuring in the last three years, and more than two-thirds reported that productivity and profits were up as result. Next on the agenda for these firms is developing a new compensation structure, and 78 percent report that they are considering a new, skills-based plan for both management and non-management employees. This coming shift, O'Neal said, "is not just isolated or a fad. It's an inexorable change." After [6 6 World War II 6], corporations adopted a "military model" creating hierarchical organisations where "the concept of defined tasks worked great," she said. But as the economy became global, customers were more demanding and problems became more complex. "Multi-layers kept management at a distance from its customers," O'Neal said. Now organisations must change to stay competitive. O'Neal says firms will place a greater emphasis on teams and team performance in giving raises. If your team does well, you'll do well. If it does n't do well, do n't expect a raise. Is this fair to an employee who can go the distance but is on a team that ca n't keep up? "That's an important question we used to ask a lot," said O'Neal. "It's not a question we ask any more. The more important question is -- 'Do we have results?' " To get those results, 27 percent of the companies surveyed plan to eliminate base pay increases in favour of cash bonuses and incentives, such as employee and team award programmes and education. Companies that can change their culture and view employees as business partners will do well, O'Neal said. For example, the survey rated some participants in the survey as "high-performing" companies, based on their return on equity above 16 percent. Those companies also said they already offer employees variable pay and business education, give their managers more control over pay decisions and celebrate employee and team success. "They respect employees more, they trust employees more and they value their employees more," O'Neal said. "The key to success at high-performance companies is engaging employees in a business partnership. It will improve a company's bottom line."

1: 3434750 United_States

2: 645042 New_York_City

3: 3434750 United_States

4: 8706816 Towers_Perrin

5: 8706816 Towers_Perrin

6: 32927 World_War_II

1: 3434750 United_States

[pem_pos=0, final_scores=0.4975298,
final_local_scores=0.48682535, pem=-0.78483206,
similarity=22.885592]

2: 645042 New_York_City

```
[pem_pos=1, final_scores=0.38434875,
final_local_scores=0.386329, pem=-1.7906687,
similarity=25.356476]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.543259,
final_local_scores=0.5367776, pem=-0.78483206,
similarity=24.376266]
4: 8706816 Towers_Perrin
    [pem_pos=0, final_scores=0.47321117,
final_local_scores=0.56308985, pem=0.0, similarity=20.893436]
5: 8706816 Towers_Perrin
    [pem_pos=0, final_scores=0.56759846,
final_local_scores=0.6661936, pem=0.0, similarity=23.970259]
6: 32927 World_War_II
    [pem_pos=0, final_scores=0.4478587,
final_local_scores=0.48756993, pem=0.0, similarity=18.639772]
```

Sample: b'392_Belgian'

[1 1 Belgian 1] police arrest detective in [2 2 Dutroux 2] affair . Geert de Clercq NEUFCHATEAU , [3 3 Belgium 3] 1996-08-25 Investigators arrested a senior police detective on Sunday in connection with their inquiries into [4 4 Belgium 4] 's child sex scandal , Public Prosecutor Michel Bourlet said . " Georges Zicot was arrested and will be charged with truck theft , insurance fraud and document forgery , " Bourlet told a news conference . He said there had been searches at three sites on Sunday , including one at the [5 5 Charleroi 5] judicial police headquarters where Zicot worked . Zicot , 45 , is a specialist in tackling vehicle theft . [6 6 Belgian 6] media reported that he had been questioned twice in the past two years about thefts but released both times . He was promoted to chief detective earlier this year . Bourlet said two other people had also been arrested . One was Gerard Pignon , the owner of a warehouse where stolen vehicles were allegedly stored . The other was insurer Thierry Dehaan . Bourlet said the investigation into the vehicle theft ring would be added to the inquiry into the paedophile sex scandal in which five other people have already been arrested . He said the connection was through Bernard Weinstein , an accomplice of convicted child rapist [7 7 Marc Dutroux 7] -- the central figure in the paedophile scandal that has sent shockwaves across [8 8 Europe 8] . Weinstein was found dead last weekend alongside the bodies of eight-year-olds Julie Lejeune and Melissa Russo in a house belonging to Detroux , who said they starved to death earlier this year , nine months after being abducted in June 1995 . Two other girls have been rescued and police are hunting for at least two more who [9 9 Dutroux 9] has admitted kidnapping a year ago . " [10 10 Dutroux 10] has admitted killing Weinstein after a disagreement between the accomplices in an affair of truck theft , " Bourlet said . Another four people were also questioned at the weekend but

had not been detained , Bourlet added . Anne Thily , public prosecutor in the eastern city of [11 11 Liege 11] where Julie and Melissa lived , said this was a major case involving some 50 investigators -- including two from the [12 12 U.S. 12] [13 13 Federal Bureau of Investigation 13] .

1: 3343 Belgium
2: 154694 Marc_Dutroux
3: 3343 Belgium
4: 3343 Belgium
5: 175228 Charleroi
6: 3343 Belgium
7: 154694 Marc_Dutroux
8: 9239 Europe
9: 154694 Marc_Dutroux
10: 154694 Marc_Dutroux
11: 28341707 Liège
12: 3434750 United_States
13: 11127 Federal_Bureau_of_Investigation
1: 3343 Belgium
[pem_pos=0, final_scores=0.7666177,
final_local_scores=0.76026106, pem=-0.12403572,
similarity=27.451944]
2: 154694 Marc_Dutroux
[pem_pos=0, final_scores=0.65366834,
final_local_scores=0.6903976, pem=0.0, similarity=24.692554]
3: 3343 Belgium
[pem_pos=0, final_scores=0.65360874,
final_local_scores=0.63681585, pem=-0.24180195,
similarity=24.408525]
4: 3343 Belgium
[pem_pos=0, final_scores=0.5881209,
final_local_scores=0.5652803, pem=-0.24180195,
similarity=22.273762]
5: 175228 Charleroi
[pem_pos=0, final_scores=0.42747563,
final_local_scores=0.44757873, pem=-0.2544922,
similarity=18.830322]
6: 3343 Belgium
[pem_pos=0, final_scores=0.63741845,
final_local_scores=0.6191304, pem=-0.12403572,
similarity=23.240324]
7: 154694 Marc_Dutroux
[pem_pos=0, final_scores=0.77072257,
final_local_scores=0.8182616, pem=0.0, similarity=28.508272]
8: 9239 Europe
[pem_pos=0, final_scores=0.61861604,
final_local_scores=0.6765358, pem=-0.086338006,
similarity=24.748411]
9: 154694 Marc_Dutroux
[pem_pos=0, final_scores=0.6087242,
final_local_scores=0.64130294, pem=0.0, similarity=23.227472]
10: 154694 Marc_Dutroux

```
[pem_pos=0, final_scores=0.5859577,
final_local_scores=0.61643404, pem=0.0, similarity=22.485334]
11: 28341707 Liège
    [pem_pos=0, final_scores=0.42798048,
final_local_scores=0.4260835, pem=0.0, similarity=16.804895]
12: 3434750 United_States
    [pem_pos=0, final_scores=0.48318076,
final_local_scores=0.5086784, pem=-0.14075686,
similarity=20.035147]
13: 11127 Federal_Bureau_of_Investigation
    [pem_pos=0, final_scores=0.3140738,
final_local_scores=0.36937374, pem=0.0, similarity=15.112561]
```

Sample: b'393_French'

[1 1 French 1] 1997 budget due around September 10 - Juppe .
BREGANCON , [2 2 France 2] 1996-08-25 Prime Minister [3 3 Ala
in Juppe 3] said on Sunday the draft 1997 budget and plans
for funding the social security system would be published
around September 10 . " The texts are practically ready , "
he told reporters after a weekend of talks with President [4 4
Jacques Chirac 4] in a [5 5 Riviera 5] fortress . The budget
had been widely expected a week or so later . -- [6 6 Paris 6]
newsroom +331 4221 5452

```
1: 5843419 France
2: 5843419 France
3: 434931 Alain_Juppé
4: 39764 Jacques_Chirac
5: 266388 French_Riviera
6: 22989 Paris
1: 5843419 France
    [pem_pos=0, final_scores=0.7200933,
final_local_scores=0.70109636, pem=-0.80223274,
similarity=29.374493]
2: 5843419 France
    [pem_pos=0, final_scores=0.70728934,
final_local_scores=0.68710995, pem=-0.16240834,
similarity=25.477646]
3: 434931 Alain_Juppé
    [pem_pos=0, final_scores=0.6005344,
final_local_scores=0.6789527, pem=0.0, similarity=24.351015]
4: 39764 Jacques_Chirac
    [pem_pos=0, final_scores=0.6488327,
final_local_scores=0.676173, pem=0.0, similarity=24.268063]
5: 266388 French_Riviera
    [pem_pos=3, final_scores=0.30547988,
final_local_scores=0.30722857, pem=-1.8379819,
similarity=23.253258]
6: 22989 Paris
    [pem_pos=0, final_scores=0.6162146,
final_local_scores=0.60169923, pem=-0.115301244,
similarity=22.672644]
```

Sample: b'394_19'

19 die as bus falls in river in [1 1 Pakistani Kashmir 1] . [2 2 MUZAFFARABAD 2] , [3 3 Pakistan 3] 1996-08-25 A bus fell from a mountain road into a river in Pakistan-ruled Azad (free) Kashmir on Sunday , killing at least 19 people and injuring 11 , police said . They said dead included five refugees from the Indian-ruled part of [4 4 Kashmir 4] , where [5 5 Moslem 5] militants have waged a separatist revolt since early 1990 . The police said 14 of the 45 passengers on the bus died instantly when the vehicle fell into [6 6 Kunar 6] river from a narrow road leading from the state capital [7 7 Muzaffarabad 7] to the nearby [8 8 Pakistani 8] town of [9 9 Garhi Habibullah 9] northwest . Five people died later in hospital .

1: 2745 Azad_Kashmir
2: 1119495 Muzaffarabad
3: 23235 Pakistan
4: 17337 Kashmir
5: 6037917 Islam
6: 4918169 Kunar_River
7: 1119495 Muzaffarabad
8: 23235 Pakistan
9: 14555341 Garhi_Habibullah
1: 2745 Azad_Kashmir
[pem_pos=0, final_scores=0.41612774,
final_local_scores=0.37120438, pem=0.0, similarity=15.167192]
2: 1119495 Muzaffarabad
[pem_pos=0, final_scores=0.36575556,
final_local_scores=0.33747134, pem=0.0, similarity=14.160531]
3: 23235 Pakistan
[pem_pos=0, final_scores=0.61542827,
final_local_scores=0.6047958, pem=-0.21944217,
similarity=23.331387]
4: 17337 Kashmir
[pem_pos=0, final_scores=0.35736686,
final_local_scores=0.31690067, pem=-0.072103634,
similarity=13.938772]
5: 6037917 Islam
[pem_pos=1, final_scores=0.38194114,
final_local_scores=0.4414245, pem=-1.2150909,
similarity=23.870552]
6: 4918169 Kunar_River
[pem_pos=1, final_scores=0.4526071,
final_local_scores=0.5027942, pem=-0.77599704,
similarity=23.314087]
7: 1119495 Muzaffarabad
[pem_pos=0, final_scores=0.40995032,
final_local_scores=0.38574737, pem=0.0, similarity=15.601185]
8: 23235 Pakistan
[pem_pos=0, final_scores=0.5082878,
final_local_scores=0.48776108, pem=-0.49826378,
similarity=21.355112]

```
9: 14555341 Garhi_Habibullah
    [pem_pos=0, final_scores=0.3918156,
final_local_scores=0.47720748, pem=0.0, similarity=18.330536]
```

Sample: b'395_North'

North Afghan highway opening put off , radio says . [1 1 ISL
AMABAD 1] 1996-08-25 The planned reopening of [2 2 Afghanistan
2] 's main northern [3 3 Salang 3] highway as a result of
peace talks with an opposition alliance has been put off
until Wednesday , official [4 4 Kabul Radio 4] said on
Sunday . The embattled [5 5 Afghan 5] government said last
week that the Kabul-Salang highway would be opened on Monday
or Tuesday following talks with the Supreme Coordination
Council alliance led by Jumbish-i-Milli movement of powerful
opposition warlord General [6 6 Abdul Rashid Dostum 6] . The
radio said on Sunday the postponement of the opening had been
made due to " precautions " . It did not elaborate . The [7 7
Salang 7] highway , [8 8 Afghanistan 8] 's main route to [9 9
Central Asia 9] , has been controlled by [10 10 Dostum 10]
since he began fighting President [11 11 Burhanuddin Rabbani 1
1] 's government in [12 12 Kabul 12] in January 1994 in
alliance with [13 13 Hezb-i-Islami 13] party leader [14 14 Gu
lbuddin Hekmatyar 14] , then prime minister but rival to the
president . [15 15 Hekmatyar 15] rejoined the government as
prime minister last June under a peace pact with [16 16 Rabba
ni 16] and has since been trying to persuade other opposition
factions to follow suit . Earlier this month , Jumbish denied
a [17 17 Kabul 17] government statement that the two sides
had agreed to a ceasefire in the north .

```
1: 51153 Islamabad
2: 737 Afghanistan
3: 18037690 Salang,_Afghanistan
4: 285790 Radio_Kabul
5: 737 Afghanistan
6: 2376 Abdul_Rashid_Dostum
7: 18037690 Salang,_Afghanistan
8: 737 Afghanistan
9: 6742 Central_Asia
10: 2376 Abdul_Rashid_Dostum
11: 4611 Burhanuddin_Rabbani
12: 16826 Kabul
13: 518768 Hezbi_Islami
14: 161917 Gulbuddin_Hekmatyar
15: 161917 Gulbuddin_Hekmatyar
16: 4611 Burhanuddin_Rabbani
17: 16826 Kabul
1: 51153 Islamabad
    [pem_pos=0, final_scores=0.36216408,
final_local_scores=0.36182728, pem=0.0, similarity=14.887361]
2: 737 Afghanistan
    [pem_pos=0, final_scores=0.54894286,
final_local_scores=0.52297896, pem=-0.22286008,
```

```
similarity=20.908396]
3: 18037690 Salang,_Afghanistan
    [pem_pos=2, final_scores=0.66941786,
final_local_scores=0.8575372, pem=-1.867658,
similarity=39.83695]
4: 285790 Radio_Kabul
    [pem_pos=0, final_scores=0.48693514,
final_local_scores=0.56934077, pem=0.0, similarity=21.079977]
5: 737 Afghanistan
    [pem_pos=0, final_scores=0.5220571,
final_local_scores=0.49361032, pem=-0.869566,
similarity=23.548864]
6: 2376 Abdul_Rashid_Dostum
    [pem_pos=0, final_scores=0.5946695,
final_local_scores=0.59906375, pem=0.0, similarity=21.96697]
7: 18037690 Salang,_Afghanistan
    [pem_pos=2, final_scores=0.7441435,
final_local_scores=0.9391636, pem=-1.867658,
similarity=42.272842]
8: 737 Afghanistan
    [pem_pos=0, final_scores=0.51732737,
final_local_scores=0.4884438, pem=-0.22286008,
similarity=19.877798]
9: 6742 Central_Asia
    [pem_pos=0, final_scores=0.30534297,
final_local_scores=0.32620332, pem=0.0, similarity=13.82427]
10: 2376 Abdul_Rashid_Dostum
    [pem_pos=0, final_scores=0.4812184,
final_local_scores=0.47513562, pem=0.0, similarity=18.268707]
11: 4611 Burhanuddin_Rabbani
    [pem_pos=0, final_scores=0.57138616,
final_local_scores=0.5531994, pem=0.0, similarity=20.598288]
12: 16826 Kabul
    [pem_pos=0, final_scores=0.41186088,
final_local_scores=0.3605944, pem=0.0, similarity=14.850569]
13: 518768 Hezbi_Islami
    [pem_pos=0, final_scores=0.42508882,
final_local_scores=0.43114948, pem=0.0, similarity=16.956074]
14: 161917 Gulbuddin_Hekmatyar
    [pem_pos=0, final_scores=0.5026725,
final_local_scores=0.4811204, pem=0.0, similarity=18.447306]
15: 161917 Gulbuddin_Hekmatyar
    [pem_pos=0, final_scores=0.42892694,
final_local_scores=0.40056455, pem=0.0, similarity=16.043358]
16: 4611 Burhanuddin_Rabbani
    [pem_pos=0, final_scores=0.4480449,
final_local_scores=0.4271741, pem=0.0, similarity=16.83744]
17: 16826 Kabul
    [pem_pos=0, final_scores=0.38545352,
final_local_scores=0.33174837, pem=0.0, similarity=13.989746]
```

Sample: b'396_Students'

Students burn [1 1 Hasina 1] effigy , battle police . Anis Ahmed [2 2 DHAKA 2] 1996-08-25 Students backed by opposition parties battled police and burned an effigy of Prime Minister Sheikh [3 3 Hasina 3] during a strike in the north [4 4 Bangladeshi 4] town of [5 5 Bogra 5] on Sunday . The strikers barricaded streets , attacked the local office of the ruling [6 6 Awami League 6] , fought running battles with police and set alight hundreds of copies of the popular " [7 7 Janakantha 7] " newspaper , alleging it was pro-government . Police used batons and teargas to try to disperse students who were throwing stones and home-made bombs , witnesses said . The strike , called by the main opposition [8 8 Bangladesh Nationalist Party 8] ([9 9 BNP 9]) , to denounce the deaths of four students killed by police over the last few days , coincided with a visit to the area by [10 10 Hasina 10] . Local officials said one policeman was killed by gunshots during clashes with pro-opposition students on Thursday . [11 11 Hasina 11] told a cross section of people at the [12 12 Bogra 12] police headquarters on Sunday that the government had already suspended three police officers and ordered a judicial probe into the violent incidents . The prime minister offered financial grants to the families of those killed , ordered the best possible medical care for the injured and urged [13 13 Bogra 13] residents to call off the strike . Opposition legislators walked out of parliament in [14 14 Dhaka 14] on Sunday denouncing " unprecedeted police barbarity " against opposition students and supporters . They renewed their call for the resignation of Home (Interior) Minister Rafiqul Islam . Hundreds of police raided the [15 15 Dhaka 15] university on Sunday , arresting nearly 30 outsiders who had been living in student dormitories and seizing weapons , university officials said . Police stormed ten residence halls on the campus , flushed out people at gunpoint and searched their baggage . They seized revolvers , sawn-off rifles , shotguns and knives . The students were later allowed to return . The swoop followed the resignation of the university 's Vice-Chancellor Dr. Emajuddin Ahmed on Saturday over the deteriorating law and order situation on the campus . Authorities closed down the 28,000-student university on Wednesday following gunbattles between students and police . Police said they fought armed activists from the [16 16 BNP 16] , headed by former prime minister [17 17 Begum Khaleda Zia 17] . [18 18 Hasina 18] told police the home ministry had already given a " blanket order " to arrest terrorists and possessors of illegal firearms irrespective of their political identities . Nearly 100 students have been injured in the clashes in [19 19 Dhaka 19] and [20 20 Bogra 20] , police told reporters .

1: 623439 Sheikh_Hasina

2: 56656 Dhaka

3: 623439 Sheikh_Hasina

4: 3454 Bangladesh

```
5: 2247892 Bogra
6: 404861 Bangladesh_Awami_League
7: 9116632 Janakantha
8: 309760 Bangladesh_Nationalist_Party
9: 309760 Bangladesh_Nationalist_Party
10: 623439 Sheikh_Hasina
11: 623439 Sheikh_Hasina
12: 2247892 Bogra
13: 2247892 Bogra
14: 56656 Dhaka
15: 56656 Dhaka
16: 309760 Bangladesh_Nationalist_Party
17: 393763 Khaleda_Zia
18: 623439 Sheikh_Hasina
19: 56656 Dhaka
20: 2247892 Bogra
1: 623439 Sheikh_Hasina
    [pem_pos=0, final_scores=0.56163025,
final_local_scores=0.54163516, pem=0.0, similarity=20.253187]
2: 56656 Dhaka
    [pem_pos=0, final_scores=0.4419405,
final_local_scores=0.40752923, pem=-0.021151643,
similarity=16.366224]
3: 623439 Sheikh_Hasina
    [pem_pos=0, final_scores=0.5441379,
final_local_scores=0.52252746, pem=0.0, similarity=19.682974]
4: 3454 Bangladesh
    [pem_pos=0, final_scores=0.514205,
final_local_scores=0.5081815, pem=-0.5810555,
similarity=22.414734]
5: 2247892 Bogra
    [pem_pos=0, final_scores=0.35151583,
final_local_scores=0.33235297, pem=0.0, similarity=14.007789]
6: 404861 Bangladesh_Awami_League
    [pem_pos=0, final_scores=0.3717944,
final_local_scores=0.34149322, pem=0.0, similarity=14.280552]
7: 9116632 Janakantha
    [pem_pos=0, final_scores=0.4738273,
final_local_scores=0.60110533, pem=0.0, similarity=22.027895]
8: 309760 Bangladesh_Nationalist_Party
    [pem_pos=0, final_scores=0.637515,
final_local_scores=0.63415337, pem=0.0, similarity=23.014114]
9: 309760 Bangladesh_Nationalist_Party
    [pem_pos=1, final_scores=0.47505325,
final_local_scores=0.45668858, pem=-1.374355,
similarity=25.192167]
10: 623439 Sheikh_Hasina
    [pem_pos=0, final_scores=0.5288903,
final_local_scores=0.5058717, pem=0.0, similarity=19.185934]
11: 623439 Sheikh_Hasina
    [pem_pos=0, final_scores=0.5135077,
final_local_scores=0.48906863, pem=0.0, similarity=18.684496]
```

```
12: 2247892 Bogra
    [pem_pos=0, final_scores=0.3277951,
final_local_scores=0.30644163, pem=0.0, similarity=13.234543]
13: 2247892 Bogra
    [pem_pos=0, final_scores=0.29427773,
final_local_scores=0.269829, pem=0.0, similarity=12.141949]
14: 56656 Dhaka
    [pem_pos=0, final_scores=0.3927998,
final_local_scores=0.3538505, pem=-0.021151643,
similarity=14.764344]
15: 56656 Dhaka
    [pem_pos=0, final_scores=0.3505637,
final_local_scores=0.307714, pem=-0.021151643,
similarity=13.387538]
16: 309760 Bangladesh_Nationalist_Party
    [pem_pos=1, final_scores=0.5862189,
final_local_scores=0.5781202, pem=-1.374355,
similarity=28.815928]
17: 393763 Khaleda_Zia
    [pem_pos=0, final_scores=0.53819364,
final_local_scores=0.52252346, pem=0.0, similarity=19.682856]
18: 623439 Sheikh_Hasina
    [pem_pos=0, final_scores=0.47660738,
final_local_scores=0.44876063, pem=0.0, similarity=17.481625]
19: 56656 Dhaka
    [pem_pos=0, final_scores=0.34971505,
final_local_scores=0.30678692, pem=-0.021151643,
similarity=13.359873]
20: 2247892 Bogra
    [pem_pos=0, final_scores=0.29770023,
final_local_scores=0.2735676, pem=0.0, similarity=12.253515]
```

Sample: b'397_16'

16 die as bus crashes in [1 1 Pakistani Kashmir 1] . [2 2 MU ZAFFARABAD 2] , [3 3 Pakistan 3] 1996-08-25 At least 16 people were killed and several injured on Sunday when a bus fell from a mountain road into a ravine on a river bank in Pakistan-ruled Azad (free) Kashmir , police said . They said 14 out of 45 passengers on the bus died instantly when the vehicle fell from the narrow road while on its way from the state capital [4 4 Muzaffarabad 4] to the nearby [5 5 Pak istani 5] town of [6 6 Garhi Habibullah 6] in the northwest . Two died later in hospital .

```
1: 2745 Azad_Kashmir
2: 1119495 Muzaffarabad
3: 23235 Pakistan
4: 1119495 Muzaffarabad
5: 23235 Pakistan
6: 14555341 Garhi_Habibullah
1: 2745 Azad_Kashmir
    [pem_pos=0, final_scores=0.38370997,
final_local_scores=0.34516227, pem=0.0, similarity=14.390043]
```

```
2: 1119495 Muzaffarabad
    [pem_pos=0, final_scores=0.3785897,
final_local_scores=0.34144852, pem=0.0, similarity=14.279218]
3: 23235 Pakistan
    [pem_pos=0, final_scores=0.6276223,
final_local_scores=0.63130647, pem=-0.21944217,
similarity=24.122517]
4: 1119495 Muzaffarabad
    [pem_pos=0, final_scores=0.41661197,
final_local_scores=0.38298208, pem=0.0, similarity=15.518662]
5: 23235 Pakistan
    [pem_pos=0, final_scores=0.49936295,
final_local_scores=0.49120265, pem=-0.49826378,
similarity=21.457817]
6: 14555341 Garhi_Habibullah
    [pem_pos=0, final_scores=0.40367013,
final_local_scores=0.4743026, pem=0.0, similarity=18.243849]
```

Sample: b'398_FEATURE'

FEATURE - Fertile [1 1 Ukraine 1] faces drought of cash and rain . Irene Marushko BATKIVSHCHYNA COLLECTIVE FARM , [2 2 Ukraine 2] 1996-08-26 Four shiny new green-and-yellow [3 3 John Deere 3] combines parked at this 1,750-hectare (4,325-acre) farm in the grain-growing regions south of [4 4 Kiev 4] don't fill its chief agronomist with enthusiasm . " They did a good job this year , but they need good diesel and good engine oil , " said Ivan Odnosum . " [5 5 God 5] help us if there is a breakdown , " he said of the machinery , loaned to the farm after the [6 6 Ukrainian 6] government bought it earlier this year . The country 's grain harvest this year is forecast to fall by more than 23 percent to only 28 million tonnes and two harsh factors are to blame . A drought scoured the steppes in May and June , stunting the growing wheat . And the farming sector , making the painful transition from [7 7 Soviet 7] central planning to a market economy simply has no money . [8 8 Ukraine 8] 's black soil is so fertile that a diplomat described it as " rich enough to grow rubber boots " . But recently [9 9 Ukraine 9] has been losing its reputation as a breadbasket of [10 10 Europe 10] , earned in the years before brutal forced collectivisation under [11 11 Soviet 11] dictator [12 12 Josef Stalin 12] . " The wheat will not be of a good quality this year , " said Hryhory Borsuk , a scientist at the Mironivka Wheat Institute , run by the Ukrainian Academy of Agrarian Sciences . " The temperature on the ground reached 62 degrees [13 13 Celsius 13] (143.60 Fahrenheit) this summer . We 've never seen anything so bad , " he said in an interview in his office , unlit since the government cut off electricity because of unpaid power bills . The harvest is gathered in the dry areas , but rainfall in Western Ukraine has delayed harvesting there . While Mironivka 's scientists , some unpaid for months , carry on developing new strains of wheat resistant to [14 14 U

kraine 14 's extreme continental climate , Borsuk said the lack of cash in **[15 15 Ukraine 15]** 's agricultural sector is as bad as the drought . Collective farms and **[16 16 Ukraine 16]** 's nascent private farming sector have no money for fertiliser , no money for herbicides and pesticides , no money to repair old or buy new machinery , no money for fuel , and none to buy good seed . This year 's harvest is down from last year 's 36.5 million tonne harvest , which in turn compares with 50 million tonnes in 1990 , the year before independence . The decline is all the worse for people who recall wasteful **[17 17 Soviet 17]** times , when the **[18 18 Kremlin 18]** imported grain but priced bread so cheaply that people bought it to feed their pigs . " Agriculture is very expensive , and there are no solutions to our problems on the horizon , " said Odnosum . He said his farm , its 260 workers now readying the fields for winter wheat sowing , expected the land to yield 5.2 tonnes per hectare (2.5 acres) but ended up with 3.9 tonnes -- still better than the national average of 2.11 tonnes per hectare this year . " We did not have money for fertiliser , we 're in debt for fuel , and we 're borrowing diesel , " Odnosum said . A few kilometres (miles) down the two-lane road which passes Odnosum 's farm , where the occasional horse-drawn buggy passes by , is the 1,700-hectare (4,200-acre) Shevchenko collective farm , built next to a village still neat and tidy despite **[19 19 post-Soviet 19]** decay . Chief Accountant Natalya Soprun said the collective is strapped for cash , earlier this year bartering 220 tonnes of grain for diesel to fuel six rickety grain combines and tractors badly in need of repairs and basic maintenance . The average family earns 160-200 million karbovanets a year (\$ 800 - \$ 1,000) -- but many have not been paid in months , Soprun said . " People are working out of their own dedication . " Borsuk said the farm sector has two options : " We can sit down and cry , or we can do something . " He said the government plans to increase the sowing of the winter wheat which will be harvested next year . Planted in late summer and early autumn , the grain is **[20 20 Ukraine 20]** 's largest export item and traditionally grows better than summer wheat in **[21 21 Ukraine 21]** 's soil . Further plans , he said , called for full provision of resources like fertiliser and herbicides for 25 percent of **[22 22 Ukraine 22]** 's arable land next year , for 50 percent in 1998 and 100 percent of the land in 1999 . " If we follow this we 'll be able to sell 10 million tonnes of grain by the year 2000 , " Borkus said . The government has been trying to phase out huge subsidies to the farm sector , and in a move back to pre-Soviet days has given up to 50 hectares (124 acres) of land to some 36,000 private farmers willing to go it alone . Collective farms will last for the next few years , because most private farms now only produce enough to feed themselves with maybe a little extra to sell , said one private farmer : " In the future , maybe , but it 's better

not to hurry . " Collectivisation brought us a lot of pain . And it went too fast . Privatising too quickly can also have a negative outcome . "

1: 31750 Ukraine
2: 31750 Ukraine
3: 432397 John_Deere
4: 585629 Kiev
5: 5042765 God
6: 31750 Ukraine
7: 26779 Soviet_Union
8: 31750 Ukraine
9: 31750 Ukraine
10: 9239 Europe
11: 26779 Soviet_Union
12: 15641 Joseph_Stalin
13: 19593040 Celsius
14: 31750 Ukraine
15: 31750 Ukraine
16: 31750 Ukraine
17: 26779 Soviet_Union
18: 381075 Moscow_Kremlin
19: 1609192 Post-Soviet_states
20: 31750 Ukraine
21: 31750 Ukraine
22: 31750 Ukraine
1: 31750 Ukraine
 [pem_pos=0, final_scores=0.6936731,
final_local_scores=0.6443446, pem=-0.35407752,
similarity=25.24377]
2: 31750 Ukraine
 [pem_pos=0, final_scores=0.6378886,
final_local_scores=0.58340865, pem=-0.35407752,
similarity=23.42532]
3: 432397 John_Deere
 [pem_pos=0, final_scores=0.33727664,
final_local_scores=0.47800022, pem=0.0, similarity=18.354193]
4: 585629 Kiev
 [pem_pos=0, final_scores=0.52903295,
final_local_scores=0.49423915, pem=-0.032242518,
similarity=19.014133]
5: 5042765 God
 [pem_pos=0, final_scores=0.5150166,
final_local_scores=0.6530157, pem=-0.18657415,
similarity=24.591623]
6: 31750 Ukraine
 [pem_pos=1, final_scores=0.66439706,
final_local_scores=0.61236507, pem=-1.0458583,
similarity=28.051449]
7: 26779 Soviet_Union
 [pem_pos=1, final_scores=0.3598345,
final_local_scores=0.31110364, pem=-0.097171314,
similarity=13.902099]

8: 31750 Ukraine
[pem_pos=0, final_scores=0.6186617,
final_local_scores=0.5624062, pem=-0.35407752,
similarity=22.798563]
9: 31750 Ukraine
[pem_pos=0, final_scores=0.66824,
final_local_scores=0.6165629, pem=-0.35407752,
similarity=24.414707]
10: 9239 Europe
[pem_pos=0, final_scores=0.6362439,
final_local_scores=0.64461094, pem=-0.086338006,
similarity=23.79571]
11: 26779 Soviet_Union
[pem_pos=0, final_scores=0.39889544,
final_local_scores=0.35377184, pem=-0.097171314,
similarity=15.175403]
12: 15641 Joseph_Stalin
[pem_pos=0, final_scores=0.42009604,
final_local_scores=0.43422133, pem=0.0, similarity=17.047743]
13: 19593040 Celsius
[pem_pos=0, final_scores=0.55777127,
final_local_scores=0.7050692, pem=0.0, similarity=25.130384]
14: 31750 Ukraine
[pem_pos=0, final_scores=0.6559748,
final_local_scores=0.6031651, pem=-0.35407752,
similarity=24.01489]
15: 31750 Ukraine
[pem_pos=0, final_scores=0.6842206,
final_local_scores=0.6340193, pem=-0.35407752,
similarity=24.935642]
16: 31750 Ukraine
[pem_pos=0, final_scores=0.66160005,
final_local_scores=0.60930973, pem=-0.35407752,
similarity=24.19826]
17: 26779 Soviet_Union
[pem_pos=1, final_scores=0.39342445,
final_local_scores=0.34779558, pem=-0.097171314,
similarity=14.99706]
18: 381075 Moscow_Kremlin
[pem_pos=1, final_scores=0.35530704,
final_local_scores=0.3850779, pem=-0.76193243,
similarity=19.724712]
19: 1609192 Post-Soviet_states
[pem_pos=0, final_scores=0.42193836,
final_local_scores=0.40312666, pem=0.0, similarity=16.119816]
20: 31750 Ukraine
[pem_pos=0, final_scores=0.5830894,
final_local_scores=0.52354884, pem=-0.35407752,
similarity=21.638983]
21: 31750 Ukraine
[pem_pos=0, final_scores=0.58352154,
final_local_scores=0.5240209, pem=-0.35407752,

```
similarity=21.65307]
22: 31750 Ukraine
    [pem_pos=0, final_scores=0.628238,
final_local_scores=0.57286686, pem=-0.35407752,
similarity=23.110731]
```

Sample: b'399_Gencor'

Gencor swells profit despite setbacks . Melanie Cheary [1 1 J OHANNESBURG 1] 1996-08-26 Gencor Ltd on Monday said it had swelled its year attributable profit and streamlined operations to strengthen it for the current financial year despite a variety of divisional setbacks . Announcing the group 's results for the year ended June 30 , chairman Brian Gilbertson said : " Happily the strong improvement in financial performance is not an illusion arising from the recent weakness of the rand relative to the dollar . " Gencor raised attributable earnings to 1,803 million rand from 1,003 million rand previously - in dollar terms an increase to \$ 469 million from \$ 279 million - and won despite the group 's Impala Platinum Holdings Ltd posting dismal results . " Not everything has gone well . We 've had substantial production difficulties at a number of our operations . The most obvious one with the greatest effect on the corporation was at Impala where we had the furnace failure , " Gilbertson said . Implats posted year attributable profit of 176 million rand from 281 million previously . Not only did the company lock in 3.92 rand per dollar in February / March , but it suffered output losses due to a furnace shutdown last August . The rand was last bid at 4.5350 against the dollar . Nor was Implats the only operation to fail output targets . Ingwe Coal Corporation Ltd was hit hard by heavy rains . It forfeited nearly one million tonnes of production to flooding at its mines in Mpumulanga province . But Gilbertson said the greatest gloom in the year came from the [2 2 European Commission 2] 's blocking of Implats ' proposed merger with [3 3 Lonrho Plc 3] 's platinum interests . " The big disappointment for the year was the failure of the platinum merger . From Gencor 's perspective we are taking the position that it is not on , " Gilbertson said . Looking ahead to the current financial year , he said that Gencor would boost earnings further . " Gencor is well placed to take up the challenges of the future . The group is soundly structured and prudently financed and is blessed with an excellent portfolio of world-class businesses . I think we can look forward to further growth " . Citing the disposal of Gencor 's stake in industrial holdings group Malbak Ltd for one billion rand among other smaller disposals , Gilbertson said Gencor had pruned its portfolio to concentrate on core assets . " We 've tried to clean up our overall portfolio by disposing of non-core assets . The biggest of those was Malbak . Overall a very substantial pruning ... leaving Gencor clearly structured along commodity lines , " Gilbertson said . He

added that the group still had about 85 million rand after tax invested in various shares , which were market for disposal when required . Referring to Gencor 's increased 41.5 percent stake in Ingwe , Gilbertson said the group would wait for coal shares to be cheaper before considering snapping up more . " I think we would have liked to have more of Ingwe but it 's the question of adding value . When we bought the shares we bought ... at quite a bit lower than the price is now ... maybe if we just wait another few years for the next downturn in coal . " Finally , Gilbertson said a growing proportion of Gencor 's income was coming from offshore and unlisted investments . -- [4 4 Johannesburg 4] newsroom +27 11 482-1003

1: 58893 Johannesburg

2: 9974 European_Commission

3: 1521351 Lonmin

4: 58893 Johannesburg

1: 58893 Johannesburg

[pem_pos=0, final_scores=0.58347267, final_local_scores=0.5724663, pem=-0.09562059, similarity=21.69325]

2: 9974 European_Commission

[pem_pos=0, final_scores=0.40788174, final_local_scores=0.5108399, pem=0.0, similarity=19.334194]

3: 1521351 Lonmin

[pem_pos=1, final_scores=0.32393283, final_local_scores=0.40377975, pem=0.0, similarity=16.139305]

4: 58893 Johannesburg

[pem_pos=0, final_scores=0.5566015, final_local_scores=0.54311365, pem=-0.09562059, similarity=20.817308]

Sample: b'400_Advanced'

Advanced Medical buying IVAC Medical . [1 1 SAN DIEGO 1]
1996-08-26 Advanced Medical Inc. said Monday it will buy IVAC Medical Systems Inc. , a former Eli Lilly & Co . unit , for about \$ 400 million in cash , creating one of the world 's largest makers of intravenous infusion therapy products . Under the agreement , IVAC and Advanced Medical 's wholly owned subsidiary , IMED Corp. , will merge to form a new company that will develop and manufacture infusion pumps that regulate the amount of intravenous fluid being administered to a patient , as well as proprietary disposable products . The combined company will have estimated revenues of \$ 353 million . Advanced Medical , through IMED , is already one of the nation 's largest developers and manufacturers of intravenous infusion pumps and proprietary disposable products . It has sales in 38 foreign countries . San Diego-based IVAC is a major provider of infusion systems and related technologies to the health-care industry . It has manufacturing plants in [2 2 San Diego 2] ; [3 3 Creedmoor 3] , N.C. ; [4 4 Hampshire 4] , [5 5 England 5] ; and [6 6 Tijua

na 6] , [7 7 Mexico 7] , and distributes its products in more than 120 countries . Eli Lilly sold IVAC on Dec. 31 , 1994 to DLJ Merchant Banking Partners LP , River Medical Inc. and other investors . Advanced Medical said it expects to take an unspecified one-time charge to pay for the merger . It did not say when the charge would be taken . " The addition of IVAC is expected to contribute to financial results in the full second quarter of 1997 , " the companies said . The merger will add to both companies ' historical leadership in infusion therapy and technology-based drug delivery devices , they said . In 1968 , IVAC introduced the world 's first infusion therapy monitoring device . A year later IVAC improved its system with the addition of an IV pump that regulated the flow of liquids through positive pressure . IMED introduced the world 's first volumetric infusion pump in 1974 . IMED had profits of \$ 1.7 million , or 6 cents per fully diluted common share , on \$ 53.9 million in revenues in the first half of 1996 . Excluding the effect of a one-time restructuring charge of \$ 17.4 million , IVAC had net income of \$ 4.2 million on net sales of \$ 112.8 million for the 1996 first half . Advanced Medical reported profits of \$ 8.4 million on sales of \$ 29.2 million in the quarter ended June 30 . William J. Mercer , IVAC president and chief executive officer , will become the president and CEO of Advanced Medical . He led the transition of IVAC to a privately held company and was previously senior vice president at Mallinckrodt Group Inc . Joseph Kuhn , Advanced Medical and IMED president , will become the new company 's executive vice president and chief financial officer . The deal is expected to close by the first quarter of 1997 , subject to regulatory approval . It has already been approved by both companies ' boards .

1: 28504 San_Diego
2: 28504 San_Diego
3: 127918 Creedmoor,_North_Carolina
4: 13861 Hampshire
5: 9316 England
6: 101739 Tijuana
7: 3966054 Mexico
1: 28504 San_Diego
[pem_pos=0, final_scores=0.51685613,
final_local_scores=0.4923486, pem=-0.18777876,
similarity=19.803547]
2: 28504 San_Diego
[pem_pos=0, final_scores=0.5060554,
final_local_scores=0.48055047, pem=-0.18777876,
similarity=19.451468]
3: 127918 Creedmoor,_North_Carolina
[pem_pos=1, final_scores=0.2787149,
final_local_scores=0.35734457, pem=-0.5115422,
similarity=17.535433]
4: 13861 Hampshire

```
[pem_pos=0, final_scores=0.47430903,
final_local_scores=0.50166416, pem=-0.3150438,
similarity=20.77363]
5: 9316 England
[pem_pos=0, final_scores=0.5260772,
final_local_scores=0.5290559, pem=-0.19053856,
similarity=20.913975]
6: 101739 Tijuana
[pem_pos=0, final_scores=0.36485112,
final_local_scores=0.36900938, pem=0.0, similarity=15.101688]
7: 3966054 Mexico
[pem_pos=0, final_scores=0.6516252,
final_local_scores=0.6918568, pem=-0.18258883,
similarity=25.729046]
```

Sample: b'401_Dow'

```
[1 1 Dow 1] rises on [2 2 Philip Morris 2] , other stocks
lower . [3 3 NEW YORK 3] 1996-08-26 The [4 4 Dow Jones 4]
industrial average opened slightly higher on Monday , boosted
by [5 5 Philip Morris 5] , which gained four to 92 . Other
shares were slightly lower , mirroring bonds . The [6 6 Dow 6]
was up eight to 5,731 , while the NASDAQ Index was off
fractionally to 1,143 and the S&P Index down one to 666 . [7 7
New York Stock Exchange 7] advances lagged declines by
476/698 while [8 8 NASDAQ 8] advances led declines 837/763 .
The 30-year [9 9 U.S. Treasury 9] bond was off 2/32 to yield
6.96 percent .
```

```
1: 47361 Dow_Jones_Industrial_Average
2: 2955912 Philip_Morris_International
3: 645042 New_York_City
4: 47361 Dow_Jones_Industrial_Average
5: 2955912 Philip_Morris_International
6: 47361 Dow_Jones_Industrial_Average
7: 21560 New_York_Stock_Exchange
8: 21559 NASDAQ
9: 53667 United_States_Department_of_the_Treasury
1: 47361 Dow_Jones_Industrial_Average
[pem_pos=2, final_scores=0.6126874,
final_local_scores=0.6108627, pem=-1.5195237,
similarity=30.58248]
2: 2955912 Philip_Morris_International
[pem_pos=2, final_scores=0.28071004,
final_local_scores=0.3087281, pem=-0.94265556,
similarity=18.429083]
3: 645042 New_York_City
[pem_pos=1, final_scores=0.35918027,
final_local_scores=0.41902375, pem=-1.7906687,
similarity=26.332151]
4: 47361 Dow_Jones_Industrial_Average
[pem_pos=1, final_scores=0.44406754,
final_local_scores=0.4266711, pem=-0.42373553,
similarity=19.126768]
```

```
5: 2955912 Philip_Morris_International
    [pem_pos=2, final_scores=0.34185314,
final_local_scores=0.37551767, pem=-0.94265556,
similarity=20.422216]
6: 47361 Dow_Jones_Industrial_Average
    [pem_pos=2, final_scores=0.5851098,
final_local_scores=0.58073837, pem=-1.5195237,
similarity=29.68351]
7: 21560 New_York_Stock_Exchange
    [pem_pos=0, final_scores=0.52499956,
final_local_scores=0.5084633, pem=0.0, similarity=19.263273]
8: 21559 NASDAQ
    [pem_pos=0, final_scores=0.43685573,
final_local_scores=0.4459287, pem=0.0, similarity=17.397114]
9: 53667 United_States_Department_of_the_Treasury
    [pem_pos=0, final_scores=0.3893807,
final_local_scores=0.4367193, pem=0.0, similarity=17.122288]
```

Sample: b'402_RUGBY'

[1 1 RUGBY LEAGUE 1] - [2 2 ST HELENS 2] CLINCH SUPER LEAGUE TITLE . [3 3 ST HELENS 3] , [4 4 England 4] 1996-08-26 [5 5 St Helens 5] completed their first league and [6 6 Challenge Cup 6] double in 30 years on Monday when they thrashed [7 7 Warrington 7] 66-14 to clinch the inaugural [8 8 Super League 8] title . [9 9 St Helens 9] secured the two points they needed in the last game of the season at [10 10 Knowsley Road 10] to win their first championship since 1975 and their first double since 1966 . In rain-soaked conditions , centre [11 11 Alan Hunte 11] grabbed a hat-trick of tries , while [12 12 Tommy Martyn 12] , [13 13 Anthony Sullivan 13] and [14 14 Paul Newlove 14] each scored two . Captain and goalkicker [15 15 Robbie Goulding 15] scored 18 points . [16 16 St Helens 16] 's triumph marked the end of [17 17 Wigan 17] 's seven-year reign as 1 British champions . [18 18 St Helens 18] needed to win on Monday to take the title -- a defeat or draw would have allowed [19 19 Wigan 19] their eighth consecutive championship . They were also the toast of [1 London Broncos 1] , who managed to scrape into the top four ahead of [20 20 Warrington 20] and qualify for the end-of-season play-offs . [21 21 Bradford 21] finished third . [22 22 St Helens 22] have now set their sights on taking the treble by winning the end-of-season premiership which begins with next Sunday 's semifinal against 2 London .

```
1: 25735 Rugby_league
2: 1095032 St_Helens_RLFC
3: 1103627 St_Helens,_Merseyside
4: 9316 England
5: 1095032 St_Helens_RLFC
6: 21050680 Challenge_Cup
7: 852392 Warrington_Wolves
8: 22751937 Super_League
9: 1095032 St_Helens_RLFC
```

```
10: 4480036 Knowsley_Road
11: 17339451 Alan_Hunte
12: 12581580 Tommy_Martyn
13: 12583751 Anthony_Sullivan
14: 12515088 Paul_Newlove
15: 3662052 Bobbie_Goulding
16: 1095032 St_Helens_RLFC
17: 1095029 Wigan_Warriors
18: 1095032 St_Helens_RLFC
19: 1095029 Wigan_Warriors
20: 852392 Warrington_Wolves
21: 1095011 Bradford_Bulls
22: 1095032 St_Helens_RLFC
1: 9316 England
    [winner: 31717 United_Kingdom, pem_pos=0,
final_scores=-0.059879422, final_local_scores=-0.0025449097,
pem=-0.48936236, similarity=6.6749964]
    gt: 9316 England, pem_pos=8, final_scores=-0.19721776,
final_local_scores=-0.20257825, pem=-4.320788,
similarity=21.54149 ]
2: 629563 London_Wasps
    [winner: 39127 The_Times, pem_pos=2,
final_scores=-0.33996665, final_local_scores=-0.29781908,
pem=-0.6931472, similarity=-1.0283606]
    RECALL MISS: 629563 London_Wasps]
1: 25735 Rugby_league
    [pem_pos=0, final_scores=0.35394126,
final_local_scores=0.33973116, pem=0.0, similarity=14.227968]
2: 1095032 St_Helens_RLFC
    [pem_pos=0, final_scores=0.5267799,
final_local_scores=0.46022564, pem=-1.1285291,
similarity=23.96088]
3: 1103627 St_Helens,_Merseyside
    [pem_pos=1, final_scores=0.341873,
final_local_scores=0.35406923, pem=-1.3087852,
similarity=21.773224]
4: 9316 England
    [pem_pos=0, final_scores=0.5019086,
final_local_scores=0.5732103, pem=-0.19053856,
similarity=22.23163]
5: 1095032 St_Helens_RLFC
    [pem_pos=0, final_scores=0.63011295,
final_local_scores=0.57310134, pem=-1.1285291,
similarity=27.329317]
6: 21050680 Challenge_Cup
    [pem_pos=0, final_scores=0.38978714,
final_local_scores=0.34579363, pem=-0.38845974,
similarity=16.521389]
7: 852392 Warrington_Wolves
    [pem_pos=1, final_scores=0.66018003,
final_local_scores=0.6121283, pem=-1.6982093,
similarity=31.591969]
```

```
8: 22751937 Super_League
    [pem_pos=0, final_scores=0.42271256,
final_local_scores=0.37527496, pem=-0.31898427,
similarity=17.023352]
9: 1095032 St_Helens_RLFC
    [pem_pos=0, final_scores=0.7169642,
final_local_scores=0.6679731, pem=-1.1285291,
similarity=30.160477]
10: 4480036 Knowsley_Road
    [pem_pos=0, final_scores=0.49197894,
final_local_scores=0.47702724, pem=0.0, similarity=18.325157]
11: 17339451 Alan_Hunte
    [pem_pos=0, final_scores=0.44662786,
final_local_scores=0.48539412, pem=0.0, similarity=18.574842]
12: 12581580 Tommy_Martyn
    [pem_pos=0, final_scores=0.51181424,
final_local_scores=0.5480801, pem=0.0, similarity=20.445517]
13: 12583751 Anthony_Sullivan
    [pem_pos=1, final_scores=0.50432557,
final_local_scores=0.52662426, pem=-0.31086537,
similarity=21.495766]
14: 12515088 Paul_Newlove
    [pem_pos=0, final_scores=0.4745316,
final_local_scores=0.48755276, pem=0.0, similarity=18.63926]
15: 3662052 Bobbie_Goulding
    [pem_pos=0, final_scores=0.5748723,
final_local_scores=0.60278887, pem=0.0, similarity=22.078135]
16: 1095032 St_Helens_RLFC
    [pem_pos=0, final_scores=0.6725203,
final_local_scores=0.6194249, pem=-1.1285291,
similarity=28.711704]
17: 1095029 Wigan_Warriors
    [pem_pos=1, final_scores=0.5296822,
final_local_scores=0.4742114, pem=-1.8027086,
similarity=28.044537]
18: 1095032 St_Helens_RLFC
    [pem_pos=0, final_scores=0.7246032,
final_local_scores=0.6763175, pem=-1.1285291,
similarity=30.40949]
19: 1095029 Wigan_Warriors
    [pem_pos=1, final_scores=0.5481779,
final_local_scores=0.4944151, pem=-1.8027086,
similarity=28.647457]
20: 852392 Warrington_Wolves
    [pem_pos=1, final_scores=0.5674848,
final_local_scores=0.5108729, pem=-1.6982093,
similarity=28.570307]
21: 1095011 Bradford_Bulls
    [pem_pos=9, final_scores=0.64288753,
final_local_scores=0.59940165, pem=-3.818427,
similarity=42.742252]
22: 1095032 St_Helens_RLFC
```

```
[pem_pos=0, final_scores=0.6873308,
final_local_scores=0.63560313, pem=-1.1285291,
similarity=29.194494]
1: 2774473 London_Broncos
    [pem_pos=0, final_scores=0.04446979,
final_local_scores=-0.03387031, pem=0.0, similarity=3.0789554]

Sample: b'403_RALLYING'
RALLYING - 1,000 LAKES RALLY RESULT / WORLD CHAMPIONSHIP
STANDINGS . JYVASKLYA , [1 1 Finland 1] 1996-08-26 Result of
the 1,000 Lakes Rally which ended on Monday : 1. Tommi
Makinen ( [2 2 Finland 2] ) [3 3 Mitsubishi Lancer 3] 4 hours
4 minutes 13 seconds 2. [4 4 Juha Kankkunen 4] ( [5 5 Finland
5] ) [6 6 Toyota Celica 6] 46 seconds behind 3. [7 7 Jarmo Ky
tolehto 7] ( [8 8 Finland 8] ) 1 Ford Escort 2:37 4. [9 9 Mar
cus Gronholm 9] ( [10 10 Finland 10] ) [11 11 Toyota Celica 11
] 2:42 5. [12 12 Kenneth Eriksson 12] ( [13 13 Sweden 13] ) [1
Subaru Impreza 1] 3:22 6. [14 14 Thomas Radstrom 14] ( [15 15
Sweden 15] ) [16 16 Toyota Celica 16] 4.09 7. [17 17 Sebastian
Lindholm 17] ( [18 18 Finland 18] ) [19 19 Ford Escort 19]
5:17 8. Lasse Lampi ( [20 20 Finland 20] ) [21 21 Mitsubishi L
ancer 21] 12:01 9. [22 22 Rui Madeira 22] ( [23 23 Portugal 23
] ) [24 24 Toyota Celica 24] 16:34 10. Angelo Medeghini ( [25
25 Italy 25] ) Subaru Impreza 18:28
1: 10577 Finland
2: 10577 Finland
3: 1119339 Mitsubishi_Lancer
4: 412742 Juha_Kankkunen
5: 10577 Finland
6: 172542 Toyota_Celica
7: 20730137 Jarmo_Kytölehto
8: 10577 Finland
9: 944600 Marcus_Grönholm
10: 10577 Finland
11: 172542 Toyota_Celica
12: 1411702 Kenneth_Eriksson
13: 5058739 Sweden
14: 7364013 Thomas_Rådström
15: 5058739 Sweden
16: 172542 Toyota_Celica
17: 11835420 Sebastian_Lindholm
18: 10577 Finland
19: 2227750 Ford_Escort_(Europe)
20: 10577 Finland
21: 1119339 Mitsubishi_Lancer
22: 21213756 Rui_Madeira
23: 23033 Portugal
24: 172542 Toyota_Celica
25: 14532 Italy
1: 2227750 Ford_Escort_(Europe)
    [winner: 2227750 Ford_Escort_(Europe), pem_pos=0,
final_scores=0.014571741, final_local_scores=-0.020183362,
```

```
pem=0.0, similarity=3.4874012]
    gt: 2227750 Ford_Escort_(Europe), pem_pos=0,
final_scores=0.014571741, final_local_scores=-0.020183362,
pem=0.0, similarity=3.4874012 ]
1: 10577 Finland
    [pem_pos=0, final_scores=0.75463563,
final_local_scores=0.76933193, pem=-0.2445894,
similarity=28.378227]
2: 10577 Finland
    [pem_pos=0, final_scores=0.7521754,
final_local_scores=0.76664436, pem=-0.2445894,
similarity=28.298025]
3: 1119339 Mitsubishi_Lancer
    [pem_pos=0, final_scores=0.30721176,
final_local_scores=0.34237093, pem=0.0, similarity=14.306745]
4: 412742 Jüha_Kankkunen
    [pem_pos=0, final_scores=0.64552945,
final_local_scores=0.7029688, pem=0.0, similarity=25.067703]
5: 10577 Finland
    [pem_pos=0, final_scores=0.7321292,
final_local_scores=0.74474704, pem=-0.2445894,
similarity=27.644566]
6: 172542 Toyota_Celica
    [pem_pos=0, final_scores=0.40763134,
final_local_scores=0.41470724, pem=0.0, similarity=16.465405]
7: 20730137 Jarmo_Kytölehto
    [pem_pos=0, final_scores=0.56359035,
final_local_scores=0.6986016, pem=0.0, similarity=24.937378]
8: 10577 Finland
    [pem_pos=0, final_scores=0.73931265,
final_local_scores=0.7525939, pem=-0.2445894,
similarity=27.87873]
9: 944600 Marcus_Grönholm
    [pem_pos=0, final_scores=0.79693896,
final_local_scores=0.84617203, pem=0.0, similarity=29.341175]
10: 10577 Finland
    [pem_pos=0, final_scores=0.75769764,
final_local_scores=0.77267677, pem=-0.2445894,
similarity=28.478045]
11: 172542 Toyota_Celica
    [pem_pos=0, final_scores=0.43277377,
final_local_scores=0.44217157, pem=0.0, similarity=17.284994]
12: 1411702 Kenneth_Eriksson
    [pem_pos=0, final_scores=0.58472055,
final_local_scores=0.6543482, pem=0.0, similarity=23.616768]
13: 5058739 Sweden
    [pem_pos=0, final_scores=0.7568218,
final_local_scores=0.7906923, pem=-0.2135508,
similarity=28.84687]
14: 7364013 Thomas_Rådström
    [pem_pos=0, final_scores=0.501334,
final_local_scores=0.5859912, pem=0.0, similarity=21.576859]
```

```
15: 5058739 Sweden
    [pem_pos=0, final_scores=0.77845496,
final_local_scores=0.81432325, pem=-0.2135508,
similarity=29.552065]
16: 172542 Toyota_Celica
    [pem_pos=0, final_scores=0.43805438,
final_local_scores=0.44793987, pem=0.0, similarity=17.457132]
17: 11835420 Sebastian_Lindholm
    [pem_pos=0, final_scores=0.45439708,
final_local_scores=0.5828591, pem=0.0, similarity=21.48339]
18: 10577 Finland
    [pem_pos=0, final_scores=0.7507631,
final_local_scores=0.7651018, pem=-0.2445894,
similarity=28.251991]
19: 2227750 Ford_Escort_(Europe)
    [pem_pos=0, final_scores=0.08570008,
final_local_scores=0.06652649, pem=0.0, similarity=6.074996]
20: 10577 Finland
    [pem_pos=0, final_scores=0.7713495,
final_local_scores=0.7875893, pem=-0.2445894,
similarity=28.923065]
21: 1119339 Mitsubishi_Lancer
    [pem_pos=0, final_scores=0.32444233,
final_local_scores=0.36119273, pem=0.0, similarity=14.868424]
22: 21213756 Rui_Madeira
    [pem_pos=0, final_scores=0.41279876,
final_local_scores=0.5150341, pem=0.0, similarity=19.459356]
23: 23033 Portugal
    [pem_pos=0, final_scores=0.56346166,
final_local_scores=0.6239839, pem=-0.30942732,
similarity=24.39335]
24: 172542 Toyota_Celica
    [pem_pos=0, final_scores=0.42098957,
final_local_scores=0.42929912, pem=0.0, similarity=16.900854]
25: 14532 Italy
    [pem_pos=0, final_scores=0.72288615,
final_local_scores=0.7846149, pem=-0.20612453,
similarity=28.625126]
1: 527976 Subaru_Impreza
    [pem_pos=0, final_scores=0.15202294,
final_local_scores=0.14112511, pem=0.0, similarity=8.3011675]
```

Sample: b'404_RALLYING'

RALLYING - MAKINEN STEPS UP TITLE BID WITH LAKES WIN . [1 1 J YVASKYLA 1] , [2 2 Finland 2] 1996-08-26 Tommi Makinen took a significant step towards becoming world rally champion with a brilliant victory in the 1000 Lakes Rally on Monday . [3 3 Mitsubishi 3] driver Makinen stopped experienced fellow [4 4 Finland 4] [5 5 Juha Kankkunen 5] in his tracks on the final day of the 1,452-km rally , doubling his lead on the first two decisive stages . " This was the most difficult win - three days at 125 percent effort , " said Makinen , whose success

completed his 1,000 Lakes hat-trick . [6 6 Kankkunen 6] was runner-up in his [7 7 Toyota 7] as [8 8 Finland 8] 's [9 9 Jarmo Kytolehto 9] produced a remarkable drive to finish third in his [10 10 Ford 10] . [11 11 Swede 11] [12 12 Kenneth Eriksson 12] kept [13 13 Subaru 13] in the hunt for the manufacturers ' title with fifth place in spite of a gearbox problem that nearly forced him off the road close to the end of the event . Maakinen 's and [14 14 Mitsubishi 14] 's positions were strengthened by the late retirement of [15 15 Spain 15] 's [16 16 Carlos Sainz 16] when his [17 17 Ford 17] gearbox failed . Makinen , with 95 points , now leads his nearest championship rival , [18 18 Sainz 18] , by 32 points .

1: 258746 Jyväskylä
2: 10577 Finland
3: 203872 Mitsubishi_Motors
4: 10577 Finland
5: 412742 Juha_Kankkunen
6: 412742 Juha_Kankkunen
7: 30984 Toyota
8: 10577 Finland
9: 20730137 Jarmo_Kytölehto
10: 30433662 Ford_Motor_Company
11: 5058739 Sweden
12: 1411702 Kenneth_Eriksson
13: 12743920 Subaru_World_Rally_Team
14: 203872 Mitsubishi_Motors
15: 26667 Spain
16: 407139 Carlos_Sainz
17: 30433662 Ford_Motor_Company
18: 407139 Carlos_Sainz
1: 258746 Jyväskylä
[pem_pos=0, final_scores=0.47040594,
final_local_scores=0.5446837, pem=0.0, similarity=20.34416]
2: 10577 Finland
[pem_pos=0, final_scores=0.7391215,
final_local_scores=0.7873029, pem=-0.2445894,
similarity=28.914518]
3: 203872 Mitsubishi_Motors
[pem_pos=1, final_scores=0.37107795,
final_local_scores=0.3817414, pem=-1.328055,
similarity=22.703808]
4: 10577 Finland
[pem_pos=2, final_scores=0.51116157,
final_local_scores=0.53829134, pem=-2.189793,
similarity=32.061836]
5: 412742 Juha_Kankkunen
[pem_pos=0, final_scores=0.64263684,
final_local_scores=0.6671242, pem=0.0, similarity=23.99803]
6: 412742 Juha_Kankkunen
[pem_pos=0, final_scores=0.64534646,
final_local_scores=0.67008406, pem=0.0, similarity=24.086357]
7: 30984 Toyota

```
[pem_pos=0, final_scores=0.490543,
final_local_scores=0.49369955, pem=-0.17762902,
similarity=19.788666]
8: 10577 Finland
    [pem_pos=0, final_scores=0.70863587,
final_local_scores=0.75400203, pem=-0.2445894,
similarity=27.920753]
9: 20730137 Jarmo_Kytölehto
    [pem_pos=0, final_scores=0.5929318,
final_local_scores=0.7338797, pem=0.0, similarity=25.990147]
10: 30433662 Ford_Motor_Company
    [pem_pos=0, final_scores=0.51845926,
final_local_scores=0.5218714, pem=-0.3393453,
similarity=21.508808]
11: 5058739 Sweden
    [pem_pos=0, final_scores=0.5741525,
final_local_scores=0.6240073, pem=-0.6464619,
similarity=26.226894]
12: 1411702 Kenneth_Eriksson
    [pem_pos=0, final_scores=0.5708398,
final_local_scores=0.62680054, pem=0.0, similarity=22.794691]
13: 12743920 Subaru_World_Rally_Team
    [pem_pos=1, final_scores=0.3348655,
final_local_scores=0.3217236, pem=-2.0824215,
similarity=25.01512]
14: 203872 Mitsubishi_Motors
    [pem_pos=1, final_scores=0.3508116,
final_local_scores=0.35960346, pem=-1.328055,
similarity=22.043167]
15: 26667 Spain
    [pem_pos=0, final_scores=0.5884058,
final_local_scores=0.6728056, pem=-0.18857051,
similarity=25.19305]
16: 407139 Carlos_Sainz
    [pem_pos=0, final_scores=0.39692295,
final_local_scores=0.38876796, pem=0.0, similarity=15.691324]
17: 30433662 Ford_Motor_Company
    [pem_pos=0, final_scores=0.4913283,
final_local_scores=0.49223495, pem=-0.3393453,
similarity=20.6244]
18: 407139 Carlos_Sainz
    [pem_pos=0, final_scores=0.3976707,
final_local_scores=0.40161264, pem=0.0, similarity=16.074635]
```

Sample: b'405_RALLYING'

RALLYING - MAKINEN WINS 1,000 LAKES RALLY . [1 1 JYVASKYLA 1]
, [2 2 Finland 2] 1996-08-26 Tommi Makinen of [3 3 Finland 3]
, driving a [4 4 Mitsubishi 4] , on Monday won the 1,000
Lakes Rally , sixth round of the world championship .

1: 258746 Jyväskylä
2: 10577 Finland
3: 10577 Finland

```
4: 203872 Mitsubishi_Motors
1: 258746 Jyväskylä
    [pem_pos=0, final_scores=0.52358574,
final_local_scores=0.55576855, pem=0.0, similarity=20.674955]
2: 10577 Finland
    [pem_pos=0, final_scores=0.8116804,
final_local_scores=0.8025647, pem=-0.2445894,
similarity=29.369959]
3: 10577 Finland
    [pem_pos=0, final_scores=0.8056763,
final_local_scores=0.79600614, pem=-0.2445894,
similarity=29.17424]
4: 203872 Mitsubishi_Motors
    [pem_pos=1, final_scores=0.3272007,
final_local_scores=0.46659517, pem=-1.328055,
similarity=25.236013]
```

Sample: b'406_SOCCER'

SOCCER - [1 1 SHARPE 1] HITS WINNER TO EASE PRESSURE ON [2 2 LEEDS 2] . [3 3 LEEDS 3] , [4 4 England 4] 1996-08-26 Winger [5 5 Lee Sharpe 5] hit a superb strike from the edge of the penalty area to give [6 6 Leeds 6] their first win of the season on Monday and leave hapless [7 7 Wimbledon 7] anchored at the bottom of the [8 8 England 8] premier league . [9 9 Sharpe 9] repaid a huge slice of the 4.5 million pound (\$ 6.98 million) fee [10 10 Leeds 10] handed [11 11 Manchester United 11] for his services with a top-draw second-half goal to hand [12 12 Wimbledon 12] their third successive defeat . [13 13 Ian Rush 13] , the [14 14 Welsh 14] striker signed from [15 15 Liverpool 15] in the close season , set up the goal , feeding [16 16 Sharpe 16] as he galloped forward and the former [17 17 England 17] winger cut inside onto his unfavoured right foot to arc a shot into the right-hand corner of the net . The only goal of the match also brought some relief for under-fire [18 18 Leeds 18] manager [19 19 Howard Wilkinson 19] following the team 's poor start to the season . Home fans frequently booed their own side until [20 20 Sharpe 20] turned the jeers to cheers .

```
1: 23166492 Lee_Sharpe
2: 7609513 Leeds_United_A.F.C.
3: 8262427 Leeds
4: 9316 England
5: 23166492 Lee_Sharpe
6: 7609513 Leeds_United_A.F.C.
7: 740757 Wimbledon_F.C.
8: 9316 England
9: 23166492 Lee_Sharpe
10: 7609513 Leeds_United_A.F.C.
11: 19961 Manchester_United_F.C.
12: 740757 Wimbledon_F.C.
13: 413277 Ian_Rush
14: 69894 Wales
```

```
15: 18119 Liverpool_F.C.
16: 23166492 Lee_Sharpe
17: 9904 England_national_football_team
18: 7609513 Leeds_United_A.F.C.
19: 760599 Howard_Wilkinson
20: 23166492 Lee_Sharpe
1: 23166492 Lee_Sharpe
    [pem_pos=0, final_scores=1.1003493,
final_local_scores=1.1159284, pem=0.0, similarity=37.391243]
2: 7609513 Leeds_United_A.F.C.
    [pem_pos=4, final_scores=0.52314216,
final_local_scores=0.47616, pem=-3.421221,
similarity=36.904408]
3: 8262427 Leeds
    [pem_pos=0, final_scores=0.43857962,
final_local_scores=0.46803933, pem=-0.25452003,
similarity=19.44106]
4: 9316 England
    [pem_pos=0, final_scores=0.5796681,
final_local_scores=0.63577825, pem=-0.19053856,
similarity=24.098782]
5: 23166492 Lee_Sharpe
    [pem_pos=0, final_scores=0.8478227,
final_local_scores=0.8361753, pem=0.0, similarity=29.042852]
6: 7609513 Leeds_United_A.F.C.
    [pem_pos=4, final_scores=0.6152739,
final_local_scores=0.5767999, pem=-3.421221,
similarity=39.907703]
7: 740757 Wimbledon_F.C.
    [pem_pos=1, final_scores=0.642837,
final_local_scores=0.63183063, pem=-1.8899087,
similarity=33.222416]
8: 9316 England
    [pem_pos=0, final_scores=0.5186109,
final_local_scores=0.56908256, pem=-0.19053856,
similarity=22.10845]
9: 23166492 Lee_Sharpe
    [pem_pos=0, final_scores=1.3690642,
final_local_scores=1.4094586, pem=0.0, similarity=46.150772]
10: 7609513 Leeds_United_A.F.C.
    [pem_pos=4, final_scores=0.6221598,
final_local_scores=0.5843217, pem=-3.421221,
similarity=40.132168]
11: 19961 Manchester_United_F.C.
    [pem_pos=0, final_scores=0.49547493,
final_local_scores=0.44603318, pem=0.0, similarity=17.400232]
12: 740757 Wimbledon_F.C.
    [pem_pos=1, final_scores=0.65464467,
final_local_scores=0.6447287, pem=-1.8899087,
similarity=33.60732]
13: 413277 Ian_Rush
    [pem_pos=0, final_scores=0.7113289,
```

```
final_local_scores=0.7118598, pem=0.0, similarity=25.33303]
14: 69894 Wales
    [pem_pos=1, final_scores=0.41871017,
final_local_scores=0.48025858, pem=-1.0776843,
similarity=24.282202]
15: 18119 Liverpool_F.C.
    [pem_pos=1, final_scores=0.523745,
final_local_scores=0.4793774, pem=-1.1702383,
similarity=24.759228]
16: 23166492 Lee_Sharpe
    [pem_pos=0, final_scores=1.3756741,
final_local_scores=1.4166789, pem=0.0, similarity=46.366234]
17: 9904 England_national_football_team
    [pem_pos=1, final_scores=0.43292922,
final_local_scores=0.38899815, pem=-2.776464,
similarity=30.797037]
18: 7609513 Leeds_United_A.F.C.
    [pem_pos=4, final_scores=0.600295,
final_local_scores=0.56043774, pem=-3.421221,
similarity=39.419426]
19: 760599 Howard_Wilkinson
    [pem_pos=0, final_scores=0.54513395,
final_local_scores=0.5497574, pem=0.0, similarity=20.495571]
20: 23166492 Lee_Sharpe
    [pem_pos=0, final_scores=1.5037676,
final_local_scores=1.5566015, pem=0.0, similarity=50.54181]
```

Sample: b'407_SOCCER'

```
SOCCER - [1 1 ENGLISH 1] PREMIER LEAGUE SUMMARY . [2 2 LONDON
2] 1996-08-26 Summary of Monday 's [3 3 English 3] premier
league soccer match : [4 4 Leeds 4] 1 ( [5 5 Sharpe 5] 58th
minute ) [6 6 Wimbledon 6] 0 . Halftime 0-0 . Attendance
25,860 .
1: 9316 England
2: 17867 London
3: 9316 England
4: 7609513 Leeds_United_A.F.C.
5: 23166492 Lee_Sharpe
6: 740757 Wimbledon_F.C.
1: 9316 England
    [pem_pos=1, final_scores=0.5574264,
final_local_scores=0.5478627, pem=-1.3527626,
similarity=27.79556]
2: 17867 London
    [pem_pos=0, final_scores=0.7114231,
final_local_scores=0.7390919, pem=-0.09086737,
similarity=26.63984]
3: 9316 England
    [pem_pos=1, final_scores=0.43221968,
final_local_scores=0.4110934, pem=-1.3527626,
similarity=23.714092]
4: 7609513 Leeds_United_A.F.C.
```

```

    [pem_pos=4, final_scores=0.55804175,
final_local_scores=0.5623929, pem=-3.421221,
similarity=39.47777]
5: 23166492 Lee_Sharpe
    [pem_pos=8, final_scores=0.81989557,
final_local_scores=0.9016035, pem=-3.321075,
similarity=49.05588]
6: 740757 Wimbledon_F.C.
    [pem_pos=1, final_scores=0.66272146,
final_local_scores=0.68843514, pem=-1.8899087,
similarity=34.911606]

Sample: b'408_SOCCER'
SOCCER - [1 1 ENGLISH 1] PREMIER LEAGUE RESULT / STANDINGS .
[2 2 LONDON 2] 1996-08-26 Result of an [3 3 English 3]
premier league soccer match on Monday : [4 4 Leeds 4] 1 [5 5 W
imbledon 5] 0 Standings ( tabulated under played , won ,
drawn , lost , goals for , against , points ) : [6 6 Sheffield
Wednesday 6] 3 3 0 0 6 2 9 [7 7 Chelsea 7] 3 2 1 0 3 0 7 [8 8
Arsenal 8] 3 2 0 1 4 2 6 [9 9 Aston Villa 9] 3 2 0 1 4 2 6 [10
10 Manchester United 10] 3 1 2 0 7 4 5 [11 11 Sunderland 11]
3 1 2 0 4 1 5 [12 12 Liverpool 12] 3 1 2 0 5 3 5 [13 13 Evert
on 13] 3 1 2 0 4 2 5 [14 14 Tottenham 14] 3 1 2 0 3 1 5 [15 15
Nottingham Forest 15] 3 1 1 1 5 5 4 [16 16 Leeds 16] 3 1 1 1
4 5 4 [17 17 West Ham 17] 3 1 1 1 3 4 4 [18 18 Leicester 18]
3 1 1 1 2 3 4 [19 19 Newcastle 19] 3 1 0 2 3 4 3 [20 20 Midd
esbrough 20] 3 0 2 1 4 5 2 [21 21 Derby 21] 3 0 2 1 4 6 2 [22
22 Southampton 22] 3 0 1 2 2 4 1 [23 23 Blackburn 23] 3 0 1 2
2 5 1 [24 24 Coventry 24] 3 0 1 2 1 6 1 [25 25 Wimbledon 25]
3 0 0 3 0 6 0
1: 9316 England
2: 17867 London
3: 9316 England
4: 7609513 Leeds_United_A.F.C.
5: 740757 Wimbledon_F.C.
6: 177538 Sheffield_Wednesday_F.C.
7: 7473 Chelsea_F.C.
8: 2174 Arsenal_F.C.
9: 42173 Aston_Villa_F.C.
10: 19961 Manchester_United_F.C.
11: 184474 Sunderland_A.F.C.
12: 18119 Liverpool_F.C.
13: 91155 Everton_F.C.
14: 68198 Tottenham_Hotspur_F.C.
15: 157294 Nottingham_Forest_F.C.
16: 7609513 Leeds_United_A.F.C.
17: 46417 West_Ham_United_F.C.
18: 298602 Leicester_City_F.C.
19: 57802 Newcastle_United_F.C.
20: 298603 Middlesbrough_F.C.
21: 4721303 Derby_County_F.C.
22: 81576 Southampton_F.C.

```

23: 298599 Blackburn_Rovers_F.C.
24: 435021 Coventry_City_F.C.
25: 740757 Wimbledon_F.C.
1: 9316 England
[pem_pos=1, final_scores=0.4872523,
final_local_scores=0.5377739, pem=-1.3527626,
similarity=27.494493]
2: 17867 London
[pem_pos=0, final_scores=0.6605023,
final_local_scores=0.7355251, pem=-0.09086737,
similarity=26.533398]
3: 9316 England
[pem_pos=1, final_scores=0.5937337,
final_local_scores=0.6540887, pem=-1.3527626,
similarity=30.965555]
4: 7609513 Leeds_United_A.F.C.
[pem_pos=4, final_scores=0.65366906,
final_local_scores=0.5992858, pem=-3.421221,
similarity=40.578728]
5: 740757 Wimbledon_F.C.
[pem_pos=1, final_scores=0.7679865,
final_local_scores=0.7427282, pem=-1.8899087,
similarity=36.53182]
6: 177538 Sheffield_Wednesday_F.C.
[pem_pos=0, final_scores=0.32715696,
final_local_scores=0.27060074, pem=0.0, similarity=12.164979]
7: 7473 Chelsea_F.C.
[pem_pos=0, final_scores=0.58833855,
final_local_scores=0.5395588, pem=-0.51707125,
similarity=23.003138]
8: 2174 Arsenal_F.C.
[pem_pos=0, final_scores=0.5399353,
final_local_scores=0.47852522, pem=-0.15820743,
similarity=19.230217]
9: 42173 Aston_Villa_F.C.
[pem_pos=0, final_scores=0.4851337,
final_local_scores=0.41410434, pem=0.0, similarity=16.447412]
10: 19961 Manchester_United_F.C.
[pem_pos=0, final_scores=0.42186385,
final_local_scores=0.35271886, pem=0.0, similarity=14.615547]
11: 184474 Sunderland_A.F.C.
[pem_pos=0, final_scores=0.7156746,
final_local_scores=0.669674, pem=-0.39640337,
similarity=26.229824]
12: 18119 Liverpool_F.C.
[pem_pos=1, final_scores=0.6688809,
final_local_scores=0.6200631, pem=-1.1702383,
similarity=28.957571]
13: 91155 Everton_F.C.
[pem_pos=0, final_scores=0.51363593,
final_local_scores=0.44603485, pem=-0.17171584,
similarity=18.3341]

```
14: 68198 Tottenham_Hotspur_F.C.  
    [pem_pos=1, final_scores=0.4143772,  
final_local_scores=0.33866483, pem=-0.99701256,  
similarity=19.618057]  
15: 157294 Nottingham_Forest_F.C.  
    [pem_pos=0, final_scores=0.5013456,  
final_local_scores=0.43558586, pem=0.0, similarity=17.088463]  
16: 7609513 Leeds_United_A.F.C.  
    [pem_pos=4, final_scores=0.7652438,  
final_local_scores=0.72116417, pem=-3.421221,  
similarity=44.21582]  
17: 46417 West_Ham_United_F.C.  
    [pem_pos=0, final_scores=0.694696,  
final_local_scores=0.6448485, pem=-0.5127755,  
similarity=26.121832]  
18: 298602 Leicester_City_F.C.  
    [pem_pos=3, final_scores=0.9275978,  
final_local_scores=0.90167004, pem=-3.2133117,  
similarity=48.471836]  
19: 57802 Newcastle_United_F.C.  
    [pem_pos=2, final_scores=0.68134815,  
final_local_scores=0.6286826, pem=-2.4885974,  
similarity=36.384235]  
20: 298603 Middlesbrough_F.C.  
    [pem_pos=0, final_scores=0.6069811,  
final_local_scores=0.55089635, pem=-0.5175458,  
similarity=23.344053]  
21: 4721303 Derby_County_F.C.  
    [pem_pos=5, final_scores=0.79122335,  
final_local_scores=0.75361186, pem=-3.703934,  
similarity=46.721565]  
22: 81576 Southampton_F.C.  
    [pem_pos=1, final_scores=0.8029571,  
final_local_scores=0.7747156, pem=-1.1503892,  
similarity=33.464767]  
23: 298599 Blackburn_Rovers_F.C.  
    [pem_pos=1, final_scores=0.6553963,  
final_local_scores=0.60020393, pem=-2.734066,  
similarity=36.86927]  
24: 435021 Coventry_City_F.C.  
    [pem_pos=2, final_scores=0.65500224,  
final_local_scores=0.60481566, pem=-3.1334689,  
similarity=39.17891]  
25: 740757 Wimbledon_F.C.  
    [pem_pos=1, final_scores=0.65479165,  
final_local_scores=0.61908, pem=-1.8899087,  
similarity=32.84191]
```

Sample: b'409_CRICKET'

CRICKET - [1 1 PAKISTAN 1] 'S [2 2 WASIM AKRAM 2] JOINS 300
CLUB . [3 3 LONDON 3] 1996-08-26 [4 4 Wasim Akram 4] 's three-
wicket haul for [5 5 Pakistan 5] in [6 6 England 6] 's second

innings at [7 7 The Oval 7] on Monday gave him his 300th test match wicket . [8 8 Wasim 8] becomes the 11th player to join the 300-club of bowlers and the second [9 9 Pakistani 9] , after [10 10 Imran Khan 10] , to achieve the feat . Other cricketers who have taken over 300 Test wickets : [11 11 Kapil Dev 11] ([12 12 India 12]) 434 wickets , 131 Tests [13 13 Richard Hadlee 13] ([14 14 New Zealand 14]) 431 , 86 [15 15 Ian Botham 15] ([16 16 England 16]) 383 , 102 [17 17 Malcolm Marshall 17] ([18 18 West Indies 18]) 376 , 81 [19 19 Imran Khan 19] ([20 20 Pakistan 20]) 362 , 88 [21 21 Dennis Lillee 21] ([22 22 Australia 22]) 355 , 70 [23 23 Bob Willis 23] ([24 24 England 24]) 325 , 90 [25 25 Lance Gibbs 25] ([26 26 West Indies 26]) 309 , 79 [27 27 Fred Trueman 27] ([28 28 England 28]) 307 , 67 [29 29 Courtney Walsh 29] ([30 30 West Indies 30]) 309 , 82 [31 31 Wasim Akram 31] ([32 32 Pakistan 32]) 300 , 70

1: 23235 Pakistan
2: 501393 Wasim_Akram
3: 17867 London
4: 501393 Wasim_Akram
5: 759295 Pakistan_national_cricket_team
6: 759125 England_cricket_team
7: 387687 The_Oval
8: 501393 Wasim_Akram
9: 23235 Pakistan
10: 413184 Imran_Khan
11: 288197 Kapil_Dev
12: 14533 India
13: 66143 Richard_Hadlee
14: 4913064 New_Zealand
15: 15524 Ian_Botham
16: 9316 England
17: 897249 Malcolm_Marshall
18: 294567 British_West_Indies
19: 413184 Imran_Khan
20: 23235 Pakistan
21: 894415 Dennis_Lillee
22: 4689264 Australia
23: 1351354 Bob_Willis
24: 9316 England
25: 2579761 Lance_Gibbs
26: 294567 British_West_Indies
27: 413202 Fred_Trueman
28: 9316 England
29: 252677 Courtney_Walsh
30: 294567 British_West_Indies
31: 501393 Wasim_Akram
32: 23235 Pakistan
1: 23235 Pakistan
[pem_pos=0, final_scores=0.5418349, final_local_scores=0.56848156, pem=-0.21944217, similarity=22.247698]

2: 501393 Wasim_Akram
[pem_pos=0, final_scores=0.42423528,
final_local_scores=0.4095111, pem=0.0, similarity=16.31034]
3: 17867 London
[pem_pos=0, final_scores=0.6962638,
final_local_scores=0.74989635, pem=-0.09086737,
similarity=26.962265]
4: 501393 Wasim_Akram
[pem_pos=0, final_scores=0.6747304,
final_local_scores=0.68313897, pem=0.0, similarity=24.475943]
5: 759295 Pakistan_national_cricket_team
[pem_pos=1, final_scores=0.70795166,
final_local_scores=0.7042131, pem=-2.7556365,
similarity=40.090416]
6: 759125 England_cricket_team
[pem_pos=2, final_scores=0.68178165,
final_local_scores=0.6648758, pem=-3.3054674,
similarity=41.906578]
7: 387687 The_Oval
[pem_pos=0, final_scores=0.4979462,
final_local_scores=0.50012916, pem=0.0, similarity=19.014565]
8: 501393 Wasim_Akram
[pem_pos=0, final_scores=0.86620826,
final_local_scores=0.89808697, pem=0.0, similarity=30.89042]
9: 23235 Pakistan
[pem_pos=0, final_scores=0.4681999,
final_local_scores=0.48804647, pem=-0.49826378,
similarity=21.363628]
10: 413184 Imran_Khan
[pem_pos=0, final_scores=0.5147471,
final_local_scores=0.52739435, pem=-0.04740224,
similarity=20.085993]
11: 288197 Kapil_Dev
[pem_pos=0, final_scores=0.3821506,
final_local_scores=0.3869896, pem=0.0, similarity=15.638255]
12: 14533 India
[pem_pos=0, final_scores=0.5297976,
final_local_scores=0.5745594, pem=-0.13622136,
similarity=21.976503]
13: 66143 Richard_Hadlee
[pem_pos=0, final_scores=0.6721927,
final_local_scores=0.6896637, pem=0.0, similarity=24.670654]
14: 4913064 New_Zealand
[pem_pos=0, final_scores=0.49097794,
final_local_scores=0.5360874, pem=-0.23795462,
similarity=21.381664]
15: 15524 Ian_Botham
[pem_pos=0, final_scores=0.50032467,
final_local_scores=0.49127394, pem=0.0, similarity=18.750307]
16: 9316 England
[pem_pos=0, final_scores=0.49521607,
final_local_scores=0.5168999, pem=-0.19053856,

```
similarity=20.551214]
17: 897249 Malcolm_Marshall
    [pem_pos=0, final_scores=0.5090103,
final_local_scores=0.5278149, pem=0.0, similarity=19.840763]
18: 294567 British_West_Indies
    [pem_pos=4, final_scores=0.33492535,
final_local_scores=0.3842749, pem=-3.0585637,
similarity=32.19019]
19: 413184 Imran_Khan
    [pem_pos=0, final_scores=0.45252472,
final_local_scores=0.45942593, pem=-0.04740224,
similarity=18.05768]
20: 23235 Pakistan
    [pem_pos=0, final_scores=0.5120354,
final_local_scores=0.53593016, pem=-0.21944217,
similarity=21.276299]
21: 894415 Dennis_Lillee
    [pem_pos=0, final_scores=0.5337988,
final_local_scores=0.54951525, pem=0.0, similarity=20.488344]
22: 4689264 Australia
    [pem_pos=0, final_scores=0.58805877,
final_local_scores=0.6391716, pem=-0.17168477,
similarity=24.097517]
23: 1351354 Bob_Willis
    [pem_pos=0, final_scores=0.45026058,
final_local_scores=0.475267, pem=0.0, similarity=18.272629]
24: 9316 England
    [pem_pos=0, final_scores=0.5190278,
final_local_scores=0.5429106, pem=-0.19053856,
similarity=21.327425]
25: 2579761 Lance_Gibbs
    [pem_pos=0, final_scores=0.45569247,
final_local_scores=0.49532443, pem=0.0, similarity=18.871181]
26: 294567 British_West_Indies
    [pem_pos=4, final_scores=0.35947198,
final_local_scores=0.4110883, pem=-3.0585637,
similarity=32.990353]
27: 413202 Fred_Trueman
    [pem_pos=0, final_scores=0.56562114,
final_local_scores=0.605768, pem=0.0, similarity=22.16704]
28: 9316 England
    [pem_pos=0, final_scores=0.43931442,
final_local_scores=0.45583582, pem=-0.19053856,
similarity=18.728941]
29: 252677 Courtney_Walsh
    [pem_pos=0, final_scores=0.55006665,
final_local_scores=0.5693141, pem=0.0, similarity=21.079182]
30: 294567 British_West_Indies
    [pem_pos=4, final_scores=0.34684372,
final_local_scores=0.39729387, pem=-3.0585637,
similarity=32.5787]
31: 501393 Wasim_Akram
```

```
[pem_pos=0, final_scores=0.69494975,
final_local_scores=0.70522547, pem=0.0, similarity=25.135048]
32: 23235 Pakistan
    [pem_pos=0, final_scores=0.5499739,
final_local_scores=0.5773722, pem=-0.21944217,
similarity=22.513012]
```

Sample: b'410_CRICKET'

CRICKET - [1 1 ENGLISH COUNTY CHAMPIONSHIP 1] STANDINGS . [2 2 LONDON 2] 1996-08-26 [3 3 English County Championship 3]
cricket standings after Monday 's matches (tabulated under
played , won , lost , drawn , batting bonus points , bowling
bonus points , total points) : [4 4 Essex 4] 13 7 2 4 45 43
212 [5 5 Kent 5] 14 7 1 6 42 40 212 [6 6 Derbyshire 6] 13 7 2
4 41 43 208 [7 7 Leicestershire 7] 13 6 1 6 43 45 202 [8 8 Su
rrey 8] 13 6 1 6 37 48 199 [9 9 Yorkshire 9] 14 6 5 3 41 46
192 [10 10 Warwickshire 10] 13 6 4 3 32 43 180 [11 11 Middles
ex 11] 13 5 5 3 26 45 160 [12 12 Sussex 12] 13 5 6 2 27 43
156 [13 13 Somerset 13] 13 4 5 4 27 49 152 [14 14 Worcestersh
ire 14] 13 3 3 7 33 48 150 [15 15 Glamorgan 15] 13 4 5 4 36
32 144 [16 16 Hampshire 16] 13 3 5 5 28 46 137 [17 17 Glouces
tershire 17] 14 3 6 5 19 47 129 [18 18 Northamptonshire 18]
13 2 6 5 30 43 120 [19 19 Lancashire 19] 13 1 4 8 38 37 115 [2
0 20 Nottinghamshire 20] 13 1 6 6 34 40 108 [21 21 Durham 21]
14 0 9 5 22 50 87

1: 683599 County_Championship
2: 17867 London
3: 683599 County_Championship
4: 1622252 Essex_County_Cricket_Club
5: 1177719 Kent_County_Cricket_Club
6: 1829984 Derbyshire_County_Cricket_Club
7: 1622318 Leicestershire_County_Cricket_Club
8: 766575 Surrey_County_Cricket_Club
9: 869945 Yorkshire_County_Cricket_Club
10: 667292 Warwickshire_County_Cricket_Club
11: 822103 Middlesex_County_Cricket_Club
12: 1763691 Sussex_County_Cricket_Club
13: 1622178 Somerset_County_Cricket_Club
14: 1156428 Worcestershire_County_Cricket_Club
15: 1622320 Glamorgan_County_Cricket_Club
16: 1919370 Hampshire_County_Cricket_Club
17: 1622322 Gloucestershire_County_Cricket_Club
18: 530308 Northamptonshire_County_Cricket_Club
19: 749919 Lancashire_County_Cricket_Club
20: 667341 Nottinghamshire_County_Cricket_Club
21: 1176371 Durham_County_Cricket_Club

1: 683599 County_Championship
 [pem_pos=0, final_scores=0.41095167,
final_local_scores=0.35321665, pem=0.0, similarity=14.630403]
2: 17867 London
 [pem_pos=0, final_scores=0.58217317,
final_local_scores=0.67547745, pem=-0.09086737,

```
similarity=24.741457]
3: 683599 County_Championship
    [pem_pos=0, final_scores=0.3298819,
final_local_scores=0.26466027, pem=0.0, similarity=11.987703]
4: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.8710167,
final_local_scores=0.83956015, pem=-2.3459642,
similarity=41.90158]
5: 1177719 Kent_County_Cricket_Club
    [pem_pos=1, final_scores=0.8257597,
final_local_scores=0.7886718, pem=-2.5096242,
similarity=41.27298]
6: 1829984 Derbyshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.8125312,
final_local_scores=0.77409655, pem=-1.4828352,
similarity=35.254185]
7: 1622318 Leicestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.8297611,
final_local_scores=0.79690033, pem=-1.0465667,
similarity=33.562202]
8: 766575 Surrey_County_Cricket_Club
    [pem_pos=1, final_scores=0.71684605,
final_local_scores=0.6668411, pem=-1.7952642,
similarity=33.752506]
9: 869945 Yorkshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.69542533,
final_local_scores=0.6443283, pem=-1.7779728,
similarity=32.986645]
10: 667292 Warwickshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.75802356,
final_local_scores=0.71273327, pem=-1.1093602,
similarity=31.39197]
11: 822103 Middlesex_County_Cricket_Club
    [pem_pos=1, final_scores=0.7691117,
final_local_scores=0.7232859, pem=-1.5695794,
similarity=34.209625]
12: 1763691 Sussex_County_Cricket_Club
    [pem_pos=1, final_scores=0.62742686,
final_local_scores=0.5855393, pem=-1.5566913,
similarity=30.028904]
13: 1622178 Somerset_County_Cricket_Club
    [pem_pos=1, final_scores=0.6870695,
final_local_scores=0.63733447, pem=-2.4139037,
similarity=36.23623]
14: 1156428 Worcestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6854647,
final_local_scores=0.6350903, pem=-0.9164197,
similarity=28.025707]
15: 1622320 Glamorgan_County_Cricket_Club
    [pem_pos=1, final_scores=0.73428047,
final_local_scores=0.69187987, pem=-0.7903395,
similarity=29.034777]
```

```
16: 1919370 Hampshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.8159489,
final_local_scores=0.78094715, pem=-1.9232318,
similarity=37.853565]
17: 1622322 Gloucestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6334039,
final_local_scores=0.5778697, pem=-1.379793,
similarity=28.838028]
18: 530308 Northamptonshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.44027728,
final_local_scores=0.3726784, pem=-1.2069724,
similarity=21.774883]
19: 749919 Lancashire_County_Cricket_Club
    [pem_pos=1, final_scores=0.542556,
final_local_scores=0.47926515, pem=-1.7295884,
similarity=27.797712]
20: 667341 Nottinghamshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.5594476,
final_local_scores=0.49707603, pem=-1.1067253,
similarity=24.941998]
21: 1176371 Durham_County_Cricket_Club
    [pem_pos=2, final_scores=0.6762598,
final_local_scores=0.6389999, pem=-2.3952909,
similarity=36.184708]
```

Sample: b'411_CRICKET'

CRICKET - [1 1 ENGLISH COUNTY CHAMPIONSHIP 1] RESULTS . [2 2 LONDON 2] 1996-08-26 Results on the final day of four-day [3 3 English County Championship 3] cricket matches on Monday : At [4 4 Colchester 4] : [5 5 Essex 5] beat [6 6 Gloucestershire 6] by an innings and 64 runs . [7 7 Gloucestershire 7] 280 and 188 (J. Russell 57 , M. Lynch 50 ; N. Williams 5-43) . [8 8 Essex 8] 532-8 declared (G. Gooch 111 , R. Irani 91 , P. Prichard 88 , D. Robinson 72 ; M. Alleyne 4-80) . [9 9 Essex 9] 24 points , [10 10 Gloucestershire 10] 3 . At [11 11 Cardiff 11] : Match drawn . [12 12 Kent 12] 323-5 declared (C. Hooper 77 , D. Fulton 64 , N. Llong 63 , M. Walker 59) and second innings forfeited . [13 13 Glamorgan 13] first innings forfeited and 273-5 (H. Morris 118 , A. Cottney 70) . [14 14 Glamorgan 14] 5 points , [15 15 Kent 15] 6 . At [16 16 Northampton 16] : [17 17 Northamptonshire 17] beat [18 18 Sussex 18] by 6 wickets . [19 19 Sussex 19] 389 and 112 . [20 20 Northamptonshire 20] 361 and 142-4 . [21 21 Northamptonshire 21] 24 points , [22 22 Sussex 22] 8 . At [23 23 Trent Bridge 23] : Match abandoned as a draw - rain . [24 24 Nottinghamshire 24] 446-9 declared and 53-0 . [25 25 Surrey 25] 128-4 declared (A. Brown 56 not out) . [26 26 Nottinghamshire 26] 8 points , [27 27 Surrey 27] 7 . At [28 28 Worcester 28] : Match drawn . [29 29 Warwickshire 29] 310 and 162-4 declared . [30 30 Worcestershire 30] 205-9 declared (K. Spiring 52 ; A. Giles 3-12) and 164-4 (P. Weston 52) . [31 31 Worcestershire 31] 8 points , [32 32 Warwickshire 32] 10 .

At [33 33 Headingley 33] : Match drawn . [34 34 Yorkshire 34] 529-8 declared (C. White 181 , R. Blakey 109 not out , M. Moxon 66 , M. Vaughan 57) . [35 35 Lancashire 35] 323 (N. Fairbrother 86 , [36 36 M. Watkinson 36] 64 ; [37 37 D. Gough 37] 4-53) and 231-7 (N. Speak 77 , N. Fairbrother 55 ; [38 38 D. Gough 38] 4-48) . [39 39 Yorkshire 39] 11 points , [40 40 Lancashire 40] 8 . At [41 41 Leicester 41] : Match drawn . [42 42 Leicestershire 42] 353 . [43 43 Hampshire 43] 137 (G. Parsons 4-36) and 135-9 . [44 44 Leicestershire 44] 11 points , [45 45 Hampshire 45] 7 .

- 1: 683599 County_Championship
- 2: 17867 London
- 3: 683599 County_Championship
- 4: 264105 Colchester
- 5: 1622252 Essex_County_Cricket_Club
- 6: 1622322 Gloucestershire_County_Cricket_Club
- 7: 1622322 Gloucestershire_County_Cricket_Club
- 8: 1622252 Essex_County_Cricket_Club
- 9: 1622252 Essex_County_Cricket_Club
- 10: 1622322 Gloucestershire_County_Cricket_Club
- 11: 5882 Cardiff
- 12: 1177719 Kent_County_Cricket_Club
- 13: 1622320 Glamorgan_County_Cricket_Club
- 14: 1622320 Glamorgan_County_Cricket_Club
- 15: 1177719 Kent_County_Cricket_Club
- 16: 63734 Northampton
- 17: 530308 Northamptonshire_County_Cricket_Club
- 18: 1763691 Sussex_County_Cricket_Club
- 19: 1763691 Sussex_County_Cricket_Club
- 20: 530308 Northamptonshire_County_Cricket_Club
- 21: 530308 Northamptonshire_County_Cricket_Club
- 22: 1763691 Sussex_County_Cricket_Club
- 23: 849500 Trent_Bridge
- 24: 667341 Nottinghamshire_County_Cricket_Club
- 25: 766575 Surrey_County_Cricket_Club
- 26: 667341 Nottinghamshire_County_Cricket_Club
- 27: 766575 Surrey_County_Cricket_Club
- 28: 58681 Worcester
- 29: 667292 Warwickshire_County_Cricket_Club
- 30: 1156428 Worcestershire_County_Cricket_Club
- 31: 1156428 Worcestershire_County_Cricket_Club
- 32: 667292 Warwickshire_County_Cricket_Club
- 33: 437832 Headingley
- 34: 869945 Yorkshire_County_Cricket_Club
- 35: 749919 Lancashire_County_Cricket_Club
- 36: 2625808 Mike_Watkinson
- 37: 413178 Darren_Gough
- 38: 413178 Darren_Gough
- 39: 869945 Yorkshire_County_Cricket_Club
- 40: 749919 Lancashire_County_Cricket_Club
- 41: 51155 Leicester
- 42: 1622318 Leicestershire_County_Cricket_Club

```
43: 1919370 Hampshire_County_Cricket_Club
44: 1622318 Leicestershire_County_Cricket_Club
45: 1919370 Hampshire_County_Cricket_Club
1: 683599 County_Championship
    [pem_pos=0, final_scores=0.3005072,
final_local_scores=0.23668331, pem=0.0, similarity=11.152815]
2: 17867 London
    [pem_pos=0, final_scores=0.6672239,
final_local_scores=0.74201, pem=-0.09086737,
similarity=26.726921]
3: 683599 County_Championship
    [pem_pos=0, final_scores=0.24733387,
final_local_scores=0.17859957, pem=0.0, similarity=9.41948]
4: 264105 Colchester
    [pem_pos=0, final_scores=0.3876117,
final_local_scores=0.41120607, pem=-0.1350495,
similarity=17.095343]
5: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.9220434,
final_local_scores=0.8983641, pem=-2.3459642,
similarity=43.65641]
6: 1622322 Gloucestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6740009,
final_local_scores=0.6233397, pem=-1.379793,
similarity=30.194944]
7: 1622322 Gloucestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6470145,
final_local_scores=0.59386116, pem=-1.379793,
similarity=29.315245]
8: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.8212338,
final_local_scores=0.78824496, pem=-2.3459642,
similarity=40.37023]
9: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.763964,
final_local_scores=0.72568643, pem=-2.3459642,
similarity=38.50336]
10: 1622322 Gloucestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6068066,
final_local_scores=0.5499401, pem=-1.379793,
similarity=28.004553]
11: 5882 Cardiff
    [pem_pos=0, final_scores=0.6832445,
final_local_scores=0.7332403, pem=-0.20401932,
similarity=27.080555]
12: 1177719 Kent_County_Cricket_Club
    [pem_pos=1, final_scores=0.79565865,
final_local_scores=0.75856704, pem=-2.5096242,
similarity=40.374596]
13: 1622320 Glamorgan_County_Cricket_Club
    [pem_pos=1, final_scores=0.7444237,
final_local_scores=0.70517015, pem=-0.7903395,
```

```
similarity=29.431385]
14: 1622320 Glamorgan_County_Cricket_Club
    [pem_pos=1, final_scores=0.7003954,
final_local_scores=0.6570759, pem=-0.7903395,
similarity=27.996157]
15: 1177719 Kent_County_Cricket_Club
    [pem_pos=1, final_scores=0.80463654,
final_local_scores=0.768374, pem=-2.5096242,
similarity=40.667255]
16: 63734 Northampton
    [pem_pos=0, final_scores=0.478422,
final_local_scores=0.4864921, pem=-0.5110059,
similarity=21.386538]
17: 530308 Northamptonshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.42501128,
final_local_scores=0.35607195, pem=-1.2069724,
similarity=21.279312]
18: 1763691 Sussex_County_Cricket_Club
    [pem_pos=1, final_scores=0.64377934,
final_local_scores=0.6032508, pem=-1.5566913,
similarity=30.557451]
19: 1763691 Sussex_County_Cricket_Club
    [pem_pos=1, final_scores=0.6647081,
final_local_scores=0.6261123, pem=-1.5566913,
similarity=31.239681]
20: 530308 Northamptonshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.39647382,
final_local_scores=0.32489914, pem=-1.2069724,
similarity=20.349054]
21: 530308 Northamptonshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.45449805,
final_local_scores=0.38828176, pem=-1.2069724,
similarity=22.240519]
22: 1763691 Sussex_County_Cricket_Club
    [pem_pos=1, final_scores=0.71538967,
final_local_scores=0.68147427, pem=-1.5566913,
similarity=32.891792]
23: 849500 Trent_Bridge
    [pem_pos=0, final_scores=0.45290697,
final_local_scores=0.4252444, pem=0.0, similarity=16.779854]
24: 667341 Nottinghamshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6224628,
final_local_scores=0.5679872, pem=-1.1067253,
similarity=27.058128]
25: 766575 Surrey_County_Cricket_Club
    [pem_pos=1, final_scores=0.74903184,
final_local_scores=0.7037924, pem=-1.7952642,
similarity=34.855206]
26: 667341 Nottinghamshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.669888,
final_local_scores=0.61979204, pem=-1.1067253,
similarity=28.604088]
```

27: 766575 Surrey_County_Cricket_Club
[pem_pos=1, final_scores=0.7172152,
final_local_scores=0.6690375, pem=-1.7952642,
similarity=33.818054]
28: 58681 Worcester
[pem_pos=0, final_scores=0.49654454,
final_local_scores=0.52520025, pem=-0.59997296,
similarity=23.025482]
29: 667292 Warwickshire_County_Cricket_Club
[pem_pos=1, final_scores=0.71875554,
final_local_scores=0.67310214, pem=-1.1093602,
similarity=30.209297]
30: 1156428 Worcestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.70188653,
final_local_scores=0.658163, pem=-0.9164197,
similarity=28.714243]
31: 1156428 Worcestershire_County_Cricket_Club
[pem_pos=1, final_scores=0.72478086,
final_local_scores=0.6831716, pem=-0.9164197,
similarity=29.460548]
32: 667292 Warwickshire_County_Cricket_Club
[pem_pos=1, final_scores=0.700027,
final_local_scores=0.65264404, pem=-1.1093602,
similarity=29.598787]
33: 437832 Headingley
[pem_pos=1, final_scores=0.11284174,
final_local_scores=0.11876243, pem=-0.5704622,
similarity=10.736084]
34: 869945 Yorkshire_County_Cricket_Club
[pem_pos=1, final_scores=0.68036467,
final_local_scores=0.63028187, pem=-1.7779728,
similarity=32.567474]
35: 749919 Lancashire_County_Cricket_Club
[pem_pos=1, final_scores=0.5958794,
final_local_scores=0.5401798, pem=-1.7295884,
similarity=29.615526]
36: 2625808 Mike_Watkinson
[pem_pos=0, final_scores=0.47777563,
final_local_scores=0.5362459, pem=0.0, similarity=20.09236]
37: 413178 Darren_Gough
[pem_pos=0, final_scores=0.5728911,
final_local_scores=0.6311589, pem=0.0, similarity=22.924753]
38: 413178 Darren_Gough
[pem_pos=0, final_scores=0.45569927,
final_local_scores=0.50314456, pem=0.0, similarity=19.10455]
39: 869945 Yorkshire_County_Cricket_Club
[pem_pos=1, final_scores=0.6787403,
final_local_scores=0.62850755, pem=-1.7779728,
similarity=32.514523]
40: 749919 Lancashire_County_Cricket_Club
[pem_pos=1, final_scores=0.56994444,
final_local_scores=0.5118498, pem=-1.7295884,

```
similarity=28.770103]
41: 51155 Leicester
    [pem_pos=0, final_scores=0.48400146,
final_local_scores=0.48567313, pem=-0.2625849,
similarity=20.011145]
42: 1622318 Leicestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.78065234,
final_local_scores=0.7452399, pem=-1.0465667,
similarity=32.02055]
43: 1919370 Hampshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.67630094,
final_local_scores=0.6327365, pem=-1.9232318,
similarity=33.430664]
44: 1622318 Leicestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.738028,
final_local_scores=0.6986794, pem=-1.0465667,
similarity=30.631094]
45: 1919370 Hampshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.61861956,
final_local_scores=0.5697283, pem=-1.9232318,
similarity=31.550373]
```

Sample: b'412_CRICKET'
CRICKET - [1 1 PAKISTAN 1] BEAT [2 2 ENGLAND 2] BY NINE
WICKETS IN THIRD TEST . [3 3 LONDON 3] 1996-08-26 [4 4 Pakistan
4] beat [5 5 England 5] by nine wickets on the fifth day
of the third and final test at [6 6 The Oval 6] on Monday to
win the series 2-0 . Scores : [7 7 England 7] 326 and 242 ; [8
8 Pakistan 8] 521-8 declared and 48-1 .

```
1: 23235 Pakistan
2: 9316 England
3: 17867 London
4: 759295 Pakistan_national_cricket_team
5: 759125 England_cricket_team
6: 387687 The_Oval
7: 759125 England_cricket_team
8: 759295 Pakistan_national_cricket_team
1: 23235 Pakistan
    [pem_pos=0, final_scores=0.525217,
final_local_scores=0.5871514, pem=-0.21944217,
similarity=22.804842]
2: 9316 England
    [pem_pos=0, final_scores=0.42537344,
final_local_scores=0.43669945, pem=-0.19053856,
similarity=18.157873]
3: 17867 London
    [pem_pos=0, final_scores=0.6516492,
final_local_scores=0.6969279, pem=-0.09086737,
similarity=25.381582]
4: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.68063,
final_local_scores=0.665855, pem=-2.7556365,
```

```

similarity=38.945732]
5: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.761565,
final_local_scores=0.74132735, pem=-3.3054674,
similarity=44.188046]
6: 387687 The_Oval
    [pem_pos=0, final_scores=0.6029789,
final_local_scores=0.601644, pem=0.0, similarity=22.04397]
7: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.73770857,
final_local_scores=0.7152677, pem=-3.3054674,
similarity=43.410374]
8: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.78949124,
final_local_scores=0.78476936, pem=-2.7556365,
similarity=42.494373]

```

Sample: b'413_CRICKET'

CRICKET - [1 1 ENGLAND 1] V [2 2 PAKISTAN 2] FINAL TEST
SCOREBOARD . [3 3 LONDON 3] 1996-08-26 Scoreboard on the last day of the third and final test between [4 4 England 4] and [5 5 Pakistan 5] at the [6 6 Oval 6] on Monday : [7 7 England 7] first innings 326 (J. Crawley 106 , G. Thorpe 54 ; [8 8 Waqar Younis 8] 4-95) [9 9 Pakistan 9] first innings 521-8 declared ([10 10 Saeed Anwar 10] 176 , [11 11 Salim Malik 11] 100 not out , [12 12 Ijaz Ahmed 12] 61) [13 13 England 13] second innings (overnight 74-0) M. Atherton c [14 14 Inzamam-ul-Haq 14] b [15 15 Mushtaq Ahmed 15] 43 A. Stewart c [16 16 Asif Mujtaba 16] b [17 17 Mushtaq Ahmed 17] 54 N. Hussain lbw b [18 18 Mushtaq Ahmed 18] 51 G. Thorpe c [19 19 Wasim Akram 19] b [20 20 Mushtaq Ahmed 20] 9 J. Crawley c Aamir Sohail b [21 21 Wasim Akram 21] 19 N. Knight c and b [22 22 Mushtaq Ahmed 22] 8 C. Lewis lbw b [23 23 Waqar Younis 23] 4 D. Cork b [24 24 Mushtaq Ahmed 24] 26 R. Croft c [25 25 Ijaz Ahmed 25] b [26 26 Wasim Akram 26] 6 I. Salisbury not out 0 A. Mullally b [27 27 Wasim Akram 27] 0 Extras (b-6 lb-2 w-1 nb-13) 22 Total 242 Fall of wickets : 1-96 2-136 3-166 4-179 5-187 6-205 7-220 8-238 9-242 Bowling : [28 28 Wasim Akram 28] 15.4-1-67-3 , [29 29 Waqar Younis 29] 18-3-55-1 , [30 30 Mushtaq Ahmed 30] 37-10-78-6 , Aamir Sohail 2-1-4-0 , [31 31 Mohammad Akram 31] 10-3-30-0 [32 32 Pakistan 32] second innings [33 33 Saeed Anwar 33] c 1 Knight b [34 34 Mullally 34] 1 Aamir Sohail not out 29 [35 35 Ijaz Ahmed 35] not out 13 Extras (nb-5) 5 Total (for one wicket) 48 Fall of wicket : 1-7 Bowling : [36 36 Cork 36] 3-0-15-0 , [37 37 Mullally 37] 3-0-24-1 , [38 38 Croft 38] 0.4-0-9-0 Result : [39 39 Pakistan 39] won by 9 wickets First test : Lord 's - [40 40 Pakistan 40] won by 164 runs Second test : [41 41 Headingley 41] - Drawn [42 42 Pakistan 42] win series 2-0

1: 9316 England
2: 23235 Pakistan
3: 17867 London

```
4: 759125 England_cricket_team
5: 759295 Pakistan_national_cricket_team
6: 387687 The_Oval
7: 759125 England_cricket_team
8: 894433 Waqar_Younis
9: 759295 Pakistan_national_cricket_team
10: 894437 Saeed_Anwar
11: 938861 Saleem_Malik
12: 938863 Ijaz_Ahmed_(cricketer)
13: 759125 England_cricket_team
14: 896403 Inzamam-ul-Haq
15: 896405 Mushtaq_Ahmed
16: 3976606 Asif_Mujtaba
17: 896405 Mushtaq_Ahmed
18: 896405 Mushtaq_Ahmed
19: 501393 Wasim_Akram
20: 896405 Mushtaq_Ahmed
21: 501393 Wasim_Akram
22: 896405 Mushtaq_Ahmed
23: 894433 Waqar_Younis
24: 896405 Mushtaq_Ahmed
25: 938863 Ijaz_Ahmed_(cricketer)
26: 501393 Wasim_Akram
27: 501393 Wasim_Akram
28: 501393 Wasim_Akram
29: 894433 Waqar_Younis
30: 896405 Mushtaq_Ahmed
31: 1958830 Mohammad_Akram
32: 759295 Pakistan_national_cricket_team
33: 894437 Saeed_Anwar
34: 2222543 Alan_Mullally
35: 938863 Ijaz_Ahmed_(cricketer)
36: 2036840 Dominic_Cork
37: 2222543 Alan_Mullally
38: 1160680 Robert_Croft
39: 759295 Pakistan_national_cricket_team
40: 759295 Pakistan_national_cricket_team
41: 1876060 Headingley_Stadium
42: 759295 Pakistan_national_cricket_team
1: 994718 Nick_Knight
    [winner: 2026862 Ray_Knight, pem_pos=4,
final_scores=-0.5193214, final_local_scores=-0.4668001,
pem=-3.8861384, similarity=11.292892]
    RECALL MISS: 994718 Nick_Knight]
1: 9316 England
    [pem_pos=0, final_scores=0.46252382,
final_local_scores=0.562016, pem=-0.19053856,
similarity=21.89757]
2: 23235 Pakistan
    [pem_pos=0, final_scores=0.34326375,
final_local_scores=0.4209398, pem=-0.21944217,
similarity=17.844757]
```

```
3: 17867 London
    [pem_pos=0, final_scores=0.5729199,
final_local_scores=0.6940967, pem=-0.09086737,
similarity=25.297092]
4: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.81087214,
final_local_scores=0.7976524, pem=-3.3054674,
similarity=45.868896]
5: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.8676581,
final_local_scores=0.8504345, pem=-2.7556365,
similarity=44.453953]
6: 387687 The_Oval
    [pem_pos=3, final_scores=0.36518866,
final_local_scores=0.34608293, pem=-2.1624494,
similarity=26.177254]
7: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.7977863,
final_local_scores=0.78335804, pem=-3.3054674,
similarity=45.442326]
8: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.703479,
final_local_scores=0.69428355, pem=0.0, similarity=24.80852]
9: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.84974533,
final_local_scores=0.8308676, pem=-2.7556365,
similarity=43.870037]
10: 894437 Saeed_Anwar
    [pem_pos=0, final_scores=0.7684036,
final_local_scores=0.7704627, pem=0.0, similarity=27.081856]
11: 938861 Saleem_Malik
    [pem_pos=0, final_scores=0.69410247,
final_local_scores=0.7147455, pem=0.0, similarity=25.419146]
12: 938863 Ijaz_Ahmed_(cricketer)
    [pem_pos=0, final_scores=0.6174177,
final_local_scores=0.60524195, pem=0.0, similarity=22.15134]
13: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.750778,
final_local_scores=0.73200864, pem=-3.3054674,
similarity=43.909958]
14: 896403 Inzamam-ul-Haq
    [pem_pos=0, final_scores=1.044498,
final_local_scores=1.0656524, pem=0.0, similarity=35.890903]
15: 896405 Mushtaq_Ahmed
    [pem_pos=0, final_scores=0.5697371,
final_local_scores=0.5482043, pem=0.0, similarity=20.449223]
16: 3976606 Asif_Mujtaba
    [pem_pos=0, final_scores=0.81433946,
final_local_scores=0.85801023, pem=0.0, similarity=29.69445]
17: 896405 Mushtaq_Ahmed
    [pem_pos=0, final_scores=0.60621345,
final_local_scores=0.5880492, pem=0.0, similarity=21.638273]
```

18: 896405 Mushtaq_Ahmed
[pem_pos=0, final_scores=0.61719364,
final_local_scores=0.60004336, pem=0.0, similarity=21.996204]
19: 501393 Wasim_Akram
[pem_pos=0, final_scores=0.74418515,
final_local_scores=0.72768486, pem=0.0, similarity=25.80528]
20: 896405 Mushtaq_Ahmed
[pem_pos=0, final_scores=0.6133687,
final_local_scores=0.5958652, pem=0.0, similarity=21.87152]
21: 501393 Wasim_Akram
[pem_pos=0, final_scores=0.68699986,
final_local_scores=0.6652186, pem=0.0, similarity=23.941162]
22: 896405 Mushtaq_Ahmed
[pem_pos=0, final_scores=0.602359,
final_local_scores=0.58383876, pem=0.0, similarity=21.512627]
23: 894433 Waqar_Younis
[pem_pos=0, final_scores=0.699429,
final_local_scores=0.6898595, pem=0.0, similarity=24.676497]
24: 896405 Mushtaq_Ahmed
[pem_pos=0, final_scores=0.6128435,
final_local_scores=0.5952915, pem=0.0, similarity=21.854399]
25: 938863 Ijaz_Ahmed_(cricketer)
[pem_pos=0, final_scores=0.62926227,
final_local_scores=0.61818033, pem=0.0, similarity=22.537447]
26: 501393 Wasim_Akram
[pem_pos=0, final_scores=0.70266205,
final_local_scores=0.6823271, pem=0.0, similarity=24.451714]
27: 501393 Wasim_Akram
[pem_pos=0, final_scores=0.68184096,
final_local_scores=0.6595832, pem=0.0, similarity=23.772991]
28: 501393 Wasim_Akram
[pem_pos=0, final_scores=0.8049821,
final_local_scores=0.7940963, pem=0.0, similarity=27.78713]
29: 894433 Waqar_Younis
[pem_pos=0, final_scores=0.7453255,
final_local_scores=0.73999447, pem=0.0, similarity=26.172625]
30: 896405 Mushtaq_Ahmed
[pem_pos=0, final_scores=0.57914865,
final_local_scores=0.558485, pem=0.0, similarity=20.756018]
31: 1958830 Mohammad_Akram
[pem_pos=0, final_scores=0.5647885,
final_local_scores=0.58937216, pem=0.0, similarity=21.677753]
32: 759295 Pakistan_national_cricket_team
[pem_pos=1, final_scores=0.8238066,
final_local_scores=0.80253345, pem=-2.7556365,
similarity=43.02449]
33: 894437 Saeed_Anwar
[pem_pos=0, final_scores=0.79919857,
final_local_scores=0.80410147, pem=0.0, similarity=28.085705]
34: 2222543 Alan_Mullally
[pem_pos=1, final_scores=0.7475048,
final_local_scores=0.78196883, pem=-0.73930055,

```
similarity=31.445656]
35: 938863 Ijaz_Ahmed_(cricketer)
    [pem_pos=0, final_scores=0.6577876,
final_local_scores=0.6493399, pem=0.0, similarity=23.467312]
36: 2036840 Dominic_Cork
    [pem_pos=8, final_scores=0.8671221,
final_local_scores=0.89016277, pem=-3.606134,
similarity=50.264664]
37: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.6862586,
final_local_scores=0.71506673, pem=-0.73930055,
similarity=29.449162]
38: 1160680 Robert_Croft
    [pem_pos=7, final_scores=0.7974909,
final_local_scores=0.8430096, pem=-3.16984,
similarity=46.484882]
39: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.8131406,
final_local_scores=0.7908824, pem=-2.7556365,
similarity=42.6768]
40: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7724981,
final_local_scores=0.7464866, pem=-2.7556365,
similarity=41.35194]
41: 1876060 Headingley_Stadium
    [pem_pos=0, final_scores=0.41137445,
final_local_scores=0.39531845, pem=-0.30334172,
similarity=17.536425]
42: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.71266335,
final_local_scores=0.6811263, pem=-2.7556365,
similarity=39.40146]
```

Sample: b'414_CRICKET'

```
CRICKET - [1 1 PAKISTAN 1] NEED 48 RUNS TO WIN THIRD AND
FINAL TEST . [2 2 LONDON 2] 1996-08-26 [3 3 England 3] were
dismissed for 242 in their second innings on the fifth day of
the third and final test at [4 4 The Oval 4] on Monday
leaving [5 5 Pakistan 5] requiring 48 runs to win . [6 6 Paki
stan 6] lead the series 1-0 .
1: 23235 Pakistan
2: 17867 London
3: 759125 England_cricket_team
4: 387687 The_Oval
5: 759295 Pakistan_national_cricket_team
6: 759295 Pakistan_national_cricket_team
1: 23235 Pakistan
    [pem_pos=0, final_scores=0.4916827,
final_local_scores=0.55577093, pem=-0.21944217,
similarity=21.868387]
2: 17867 London
    [pem_pos=0, final_scores=0.5855041,
```

```

final_local_scores=0.6473545, pem=-0.09086737,
similarity=23.902212]
3: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.75206214,
final_local_scores=0.7429983, pem=-3.3054674,
similarity=44.23791]
4: 387687 The_Oval
    [pem_pos=0, final_scores=0.5878809,
final_local_scores=0.59443873, pem=0.0, similarity=21.82895]
5: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7672364,
final_local_scores=0.7513233, pem=-2.7556365,
similarity=41.496277]
6: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7666865,
final_local_scores=0.7507226, pem=-2.7556365,
similarity=41.47835]

```

Sample: b'415_RUGBY'

```

[1 1 RUGBY LEAGUE 1] - EUROPEAN SUPER LEAGUE RESULT / FINALS
STANDINGS . [2 2 LONDON 2] 1996-08-26 Result of a [3 3 Europe
an Super League 3] rugby league match on Monday : [4 4 St Hel
ens 4] 66 [5 5 Warrington 5] 14 Final standings ( tabulated
under played , won , drawn , lost , points for , against ,
total points ) : [6 6 St Helens 6] 22 20 0 2 950 455 40 -
champions [7 7 Wigan 7] 22 19 1 2 902 326 39 [8 8 Bradford 8]
22 17 0 5 767 409 34 London 22 12 1 9 611 462 25 [9 9 Warring
ton 9] 22 12 0 10 569 565 24 [10 10 Halifax 10] 22 10 1 11
667 576 21 1 Sheffield 22 10 0 12 599 730 20 [11 11 Oldham 11]
22 9 1 12 473 681 19 [12 12 Castleford 12] 22 9 0 13 548 599
18 [13 13 Leeds 13] 22 6 0 16 555 745 12 2 Paris 22 3 1 18
398 795 7 [14 14 Workington 14] 22 2 1 19 325 1021 5
1: 25735 Rugby_league
2: 17867 London
3: 22751937 Super_League
4: 1095032 St_Helens_RLFC
5: 852392 Warrington_Wolves
6: 1095032 St_Helens_RLFC
7: 1095029 Wigan_Warriors
8: 1095011 Bradford_Bulls
9: 852392 Warrington_Wolves
10: 1196374 Halifax_RLFC
11: 1098719 Oldham_Roughyeds
12: 1216025 Castleford_Tigers
13: 1095024 Leeds_Rhinos
14: 3646900 Workington_Town
1: 3367769 Sheffield_Eagles
    [winner: 88885 Sheffield, pem_pos=0,
final_scores=-0.44869363, final_local_scores=-0.48247248,
pem=-0.25810194, similarity=-8.90463|
    RECALL MISS: 3367769 Sheffield_Eagles]
2: 5690457 Paris_Saint-Germain_Rugby_League

```

```
[winner: 22989 Paris, pem_pos=0,
final_scores=-0.49752563, final_local_scores=-0.4622226,
pem=-0.115301244, similarity=-9.076905]
    RECALL MISS: 5690457 Paris_Saint-Germain_Rugby_League]
1: 25735 Rugby_league
    [pem_pos=0, final_scores=0.34628206,
final_local_scores=0.32955882, pem=0.0, similarity=13.924405]
2: 17867 London
    [pem_pos=0, final_scores=0.56203234,
final_local_scores=0.67323244, pem=-0.09086737,
similarity=24.674461]
3: 22751937 Super_League
    [pem_pos=0, final_scores=0.45645845,
final_local_scores=0.41099405, pem=0.0, similarity=16.354595]
4: 1095032 St_Helens_RLFC
    [pem_pos=0, final_scores=0.69612914,
final_local_scores=0.65172654, pem=-1.1285291,
similarity=29.67565]
5: 852392 Warrington_Wolves
    [pem_pos=1, final_scores=0.7249234,
final_local_scores=0.6806988, pem=-1.6982093,
similarity=33.63825]
6: 1095032 St_Helens_RLFC
    [pem_pos=0, final_scores=0.61726934,
final_local_scores=0.5655841, pem=-1.1285291,
similarity=27.104988]
7: 1095029 Wigan_Warriors
    [pem_pos=1, final_scores=0.6203989,
final_local_scores=0.5740742, pem=-1.8027086,
similarity=31.024643]
8: 1095011 Bradford_Bulls
    [pem_pos=9, final_scores=0.72949004,
final_local_scores=0.6883177, pem=-3.818427,
similarity=45.395687]
9: 852392 Warrington_Wolves
    [pem_pos=1, final_scores=0.810824,
final_local_scores=0.77453214, pem=-1.6982093,
similarity=36.438423]
10: 1196374 Halifax_RLFC
    [pem_pos=3, final_scores=0.84833735,
final_local_scores=0.82434255, pem=-2.7129183,
similarity=43.44301]
11: 1098719 Oldham_Roughyeds
    [pem_pos=1, final_scores=0.57927996,
final_local_scores=0.5363547, pem=-2.0696447,
similarity=31.350655]
12: 1216025 Castleford_Tigers
    [pem_pos=1, final_scores=0.7019616,
final_local_scores=0.66859823, pem=-0.78848183,
similarity=28.329905]
13: 1095024 Leeds_Rhinos
    [pem_pos=1, final_scores=0.543669,
```

```
final_local_scores=0.4849884, pem=-2.7286434,  
similarity=33.401524]  
14: 3646900 Workington_Town  
    [pem_pos=3, final_scores=0.44068164,  
final_local_scores=0.39061153, pem=-1.9682475,  
similarity=26.449976]
```

Sample: b'416_SOCCER'
SOCCER - [1 1 ENGLISH 1] AND [2 2 SCOTTISH 2] LEAGUE
FIXTURES - AUG 30-SEPT 1 . [3 3 LONDON 3] 1996-08-26 [4 4 English 4] and [5 5 Scottish 5] league soccer fixtures for August 30 to September 1 : Friday , August 30 : [6 6 English 6] division one - [7 7 West Bromwich 7] v [8 8 Sheffield United 8] . [9 9 English 9] division three - [10 10 Swansea 10] v 1 Lincoln . Saturday , August 31 : [11 11 English 11] division one - [12 12 Birmingham 12] v [13 13 Barnsley 13] , 2 Bradford v [14 14 Tranmere 14] , [15 15 Grimsby 15] v [16 16 Portsmouth 16] , [17 17 Huddersfield 17] v [18 18 Crystal Palace 18] , [19 19 Manchester City 19] v [20 20 Charlton 20] , [21 21 Norwich 21] v [22 22 Wolverhampton 22] , [23 23 Oldham 23] v [24 24 Ipswich 24] , [25 25 Port Vale 25] v [26 26 Oxford 26] , [27 27 Reading 27] v [28 28 Stoke 28] , [29 29 Southend 29] v [30 30 Swindon 30] . [31 31 English 31] division two - [32 32 Blackpool 32] v [33 33 Wycombe 33] , [34 34 Bournemouth 34] v [35 35 Peterborough 35] , [36 36 Bristol Rovers 36] v [37 37 Stockport 37] , [38 38 Bury 38] v [39 39 Bristol City 39] , [40 40 Crewe 40] v [41 41 Watford 41] , [42 42 Gillingham 42] v [43 43 Chesterfield 43] , [44 44 Luton 44] v [45 45 Rotherham 45] , [46 46 Millwall 46] v [47 47 Burnley 47] , [48 48 Notts County 48] v 3 York , 4 Plymouth v [49 49 Preston 49] , [50 50 Shrewsbury 50] v [51 51 Brentford 51] , [52 52 Walsall 52] v [53 53 Wrexham 53] . [54 54 English 54] division three - [55 55 Brighton 55] v [56 56 Scunthorpe 56] , 5 Cambridge v [57 57 Cardiff 57] , [58 58 Colchester 58] v [59 59 Hereford 59] , [60 60 Doncaster 60] v [61 61 Darlington 61] , [62 62 Fulham 62] v [63 63 Carlisle 63] , [64 64 Hull 64] v [65 65 Barnet 65] , [66 66 Leyton Orient 66] v [67 67 Hartlepool 67] , [68 68 Mansfield 68] v [69 69 Rochdale 69] , [70 70 Scarborough 70] v [71 71 Northampton 71] , [72 72 Torquay 72] v 6 Exeter , [73 73 Wigan 73] v [74 74 Chester 74] . [75 75 Scottish 75] division one - [76 76 East Fife 76] v [77 77 Clydebank 77] , [78 78 Greenock Morton 78] v Falkirk , [79 79 Partick 79] v [80 80 St Mirren 80] , [81 81 St Johnstone 81] v [82 82 Airdrieonians 82] , 7 Stirling v [83 83 Dundee 83] . [84 84 Scottish 84] division two - [85 85 Ayr 85] v 8 Berwick , [86 86 Clyde 86] v [87 87 Queen of the South 87] , [88 88 Dumbarton 88] v [89 89 Brechin 89] , [90 90 Livingston 90] v 9 Hamilton , [91 91 Stenhousemuir 91] v [92 92 Stranraer 92] . [93 93 Scottish 93] division three - Albion v [94 94 Cowdenbeath 94] , [95 95 Arbroath 95] v [96 96 East Stirlingshire 96] , [97 97 Inverness 97] v [98 98 Alloa 98] , [99 99 Montrose 99] v [100 100 Ross County 100] , Queen 's Park v [101 101 Forfar 101] . Sunday ,

September 1 : [102 102 English 102] division one - [103 103 Qu
eens Park Rangers 103] v [104 104 Bolton 104] .
1: 9316 England
2: 26994 Scotland
3: 17867 London
4: 9316 England
5: 26994 Scotland
6: 9316 England
7: 33921 West_Bromwich_Albion_F.C.
8: 314933 Sheffield_United_F.C.
9: 9316 England
10: 451169 Swansea_City_A.F.C.
11: 9316 England
12: 185196 Birmingham_City_F.C.
13: 333188 Barnsley_F.C.
14: 69219 Tranmere_Rovers_F.C.
15: 451129 Grimsby_Town_F.C.
16: 564301 Portsmouth_F.C.
17: 236211 Huddersfield_Town_F.C.
18: 385313 Crystal_Palace_F.C.
19: 165813 Manchester_City_F.C.
20: 6720 Charlton_Athletic_F.C.
21: 248083 Norwich_City_F.C.
22: 217730 Wolverhampton_Wanderers_F.C.
23: 451131 Oldham_Athletic_A.F.C.
24: 352843 Ipswich_Town_F.C.
25: 203456 Port_Vale_F.C.
26: 451165 Oxford_United_F.C.
27: 168007 Reading_F.C.
28: 203434 Stoke_City_F.C.
29: 451168 Southend_United_F.C.
30: 74291 Swindon_Town_F.C.
31: 9316 England
32: 245357 Blackpool_F.C.
33: 451136 Wycombe_Wanderers_F.C.
34: 451121 A.F.C._Bournemouth
35: 451132 Peterborough_United_F.C.
36: 241161 Bristol_Rovers_F.C.
37: 451135 Stockport_County_F.C.
38: 451156 Bury_F.C.
39: 4392 Bristol_City_F.C.
40: 69438 Crewe_Alexandra_F.C.
41: 85099 Watford_F.C.
42: 165278 Gillingham_F.C.
43: 451126 Chesterfield_F.C.
44: 18367 Luton_Town_F.C.
45: 386362 Rotherham_United_F.C.
46: 435024 Millwall_F.C.
47: 376725 Burnley_F.C.
48: 448599 Notts_County_F.C.
49: 429092 Preston_North_End_F.C.
50: 451207 Shrewsbury_Town_F.C.

51: 451122 Brentford_F.C.
52: 359501 Walsall_F.C.
53: 413286 Wrexham_F.C.
54: 9316 England
55: 69455 Brighton_&_Hove_Albion_F.C.
56: 451167 Scunthorpe_United_F.C.
57: 5624949 Cardiff_City_F.C.
58: 451127 Colchester_United_F.C.
59: 451201 Hereford_United_F.C.
60: 451160 Doncaster_Rovers_F.C.
61: 451159 Darlington_F.C.
62: 11228 Fulham_F.C.
63: 451157 Carlisle_United_F.C.
64: 451358 Hull_City_A.F.C.
65: 451173 Barnet_F.C.
66: 451162 Leyton_Orient_F.C.
67: 451130 Hartlepool_United_F.C.
68: 451164 Mansfield_Town_F.C.
69: 451166 Rochdale_A.F.C.
70: 946449 Scarborough_F.C.
71: 425750 Northampton_Town_F.C.
72: 193049 Torquay_United_F.C.
73: 367030 Wigan_Athletic_F.C.
74: 451195 Chester_City_F.C.
75: 26994 Scotland
76: 451245 East_Fife_F.C.
77: 1347678 Clydebank_F.C.
78: 367890 Greenock_Morton_F.C.
79: 226415 Partick_Thistle_F.C.
80: 783089 St._Mirren_F.C.
81: 393756 St._Johnstone_F.C.
82: 1216952 Airdrieonians_F.C._(1878)
83: 306504 Dundee_F.C.
84: 26994 Scotland
85: 451240 Ayr_United_F.C.
86: 451243 Clyde_F.C.
87: 880354 Queen_of_the_South_F.C.
88: 443155 Dumbarton_F.C.
89: 451241 Brechin_City_F.C.
90: 451183 Livingston_F.C.
91: 653922 Stenhousemuir_F.C.
92: 880391 Stranraer_F.C.
93: 26994 Scotland
94: 451244 Cowdenbeath_F.C.
95: 451238 Arbroath_F.C.
96: 817230 East_Stirlingshire_F.C.
97: 743540 Inverness_Caledonian_Thistle_F.C.
98: 451237 Alloa_Athletic_F.C.
99: 880408 Montrose_F.C.
100: 880380 Ross_County_F.C.
101: 408985 Forfar_Athletic_F.C.
102: 9316 England

103: 84588 Queens_Park_Rangers_F.C.
104: 298601 Bolton_Wanderers_F.C.
1: 451163 Lincoln_City_F.C.
[winner: 62705 Lincoln_Motor_Company, pem_pos=2,
final_scores=-0.57659835, final_local_scores=-0.5214546,
pem=-2.4392905, similarity=1.793711]
RECALL MISS: 451163 Lincoln_City_F.C.]
2: 431144 Bradford_City_A.F.C.
[winner: 1095011 Bradford_Bulls, pem_pos=9,
final_scores=-0.3331496, final_local_scores=-0.3976227,
pem=-3.818427, similarity=12.9890585]
RECALL MISS: 431144 Bradford_City_A.F.C.]
3: 451170 York_City_F.C.
[winner: 992980 York_City_Knights, pem_pos=18,
final_scores=-0.6464703, final_local_scores=-0.7240468,
pem=-4.2129073, similarity=5.393157]
RECALL MISS: 451170 York_City_F.C.]
4: 23857 Plymouth_Argyle_F.C.
[winner: 23508 Plymouth, pem_pos=0,
final_scores=-0.7465751, final_local_scores=-0.8042859,
pem=-0.45987546, similarity=-17.410908]
RECALL MISS: 23857 Plymouth_Argyle_F.C.]
5: 243970 Cambridge_United_F.C.
[winner: 25978572 University_of_Cambridge, pem_pos=1,
final_scores=-0.9732412, final_local_scores=-1.0205085,
pem=-1.7952901, similarity=-16.601229]
RECALL MISS: 243970 Cambridge_United_F.C.]
6: 347780 Exeter_City_F.C.
[winner: 30871776 Exeter_Chiefs, pem_pos=19,
final_scores=-0.46086472, final_local_scores=-0.50965726,
pem=-3.6095665, similarity=8.509907]
RECALL MISS: 347780 Exeter_City_F.C.]
7: 880388 Stirling_Albion_F.C.
[winner: 61311 Stirling, pem_pos=0,
final_scores=-0.48914668, final_local_scores=-0.5201857,
pem=-0.2433456, similarity=-10.110316]
RECALL MISS: 880388 Stirling_Albion_F.C.]
8: 451242 Berwick_Rangers_F.C.
[winner: 3599108 HMS_Berwick_(65), pem_pos=14,
final_scores=-0.7907595, final_local_scores=-0.74512315,
pem=-3.3529623, similarity=0.087682694]
RECALL MISS: 451242 Berwick_Rangers_F.C.]
9: 309502 Hamilton_Academical_F.C.
[winner: 86654 Hamilton,_New_Zealand, pem_pos=1,
final_scores=-0.87285614, final_local_scores=-0.90096396,
pem=-2.40051, similarity=-9.742502]
RECALL MISS: 309502 Hamilton_Academical_F.C.]
1: 9316 England
[pem_pos=1, final_scores=0.48777276,
final_local_scores=0.53340214, pem=-1.3527626,
similarity=27.36403]
2: 26994 Scotland

```
[pem_pos=0, final_scores=0.43488485,
final_local_scores=0.48023826, pem=-0.47951394,
similarity=21.028652]
3: 17867 London
    [pem_pos=0, final_scores=0.61732846,
final_local_scores=0.6935213, pem=-0.09086737,
similarity=25.279922]
4: 9316 England
    [pem_pos=1, final_scores=0.48936075,
final_local_scores=0.53513676, pem=-1.3527626,
similarity=27.415794]
5: 26994 Scotland
    [pem_pos=0, final_scores=0.40953964,
final_local_scores=0.45255244, pem=-0.47951394,
similarity=20.202452]
6: 9316 England
    [pem_pos=1, final_scores=0.589542,
final_local_scores=0.6445696, pem=-1.3527626,
similarity=30.681488]
7: 33921 West_Bromwich_Albion_F.C.
    [pem_pos=2, final_scores=0.55916834,
final_local_scores=0.50565094, pem=-1.5002033,
similarity=27.337685]
8: 314933 Sheffield_United_F.C.
    [pem_pos=0, final_scores=0.32962865,
final_local_scores=0.2476607, pem=0.0, similarity=11.480402]
9: 9316 England
    [pem_pos=1, final_scores=0.713221,
final_local_scores=0.7796702, pem=-1.3527626,
similarity=34.713158]
10: 451169 Swansea_City_A.F.C.
    [pem_pos=3, final_scores=0.8775075,
final_local_scores=0.8540054, pem=-2.5988693,
similarity=43.707993]
11: 9316 England
    [pem_pos=1, final_scores=0.57240087,
final_local_scores=0.6258455, pem=-1.3527626,
similarity=30.122725]
12: 185196 Birmingham_City_F.C.
    [pem_pos=2, final_scores=0.9104791,
final_local_scores=0.8845894, pem=-3.102209,
similarity=47.357918]
13: 333188 Barnsley_F.C.
    [pem_pos=0, final_scores=0.4206484,
final_local_scores=0.34757027, pem=-0.7126868,
similarity=18.337605]
14: 69219 Tranmere_Rovers_F.C.
    [pem_pos=0, final_scores=0.5657526,
final_local_scores=0.5121075, pem=-0.6002322,
similarity=22.636179]
15: 451129 Grimsby_Town_F.C.
    [pem_pos=2, final_scores=0.7246073,
```

```
final_local_scores=0.6816794, pem=-2.6069987,
similarity=38.60965]
16: 564301 Portsmouth_F.C.
    [pem_pos=1, final_scores=0.6233131,
final_local_scores=0.57152134, pem=-1.2302613,
similarity=27.835402]
17: 236211 Huddersfield_Town_F.C.
    [pem_pos=3, final_scores=0.7370646,
final_local_scores=0.69407433, pem=-2.2859883,
similarity=37.233833]
18: 385313 Crystal_Palace_F.C.
    [pem_pos=0, final_scores=0.51665336,
final_local_scores=0.4570002, pem=-0.51090026,
similarity=20.505865]
19: 165813 Manchester_City_F.C.
    [pem_pos=0, final_scores=0.3056906,
final_local_scores=0.23611167, pem=0.0, similarity=11.1357565]
20: 6720 Charlton_Athletic_F.C.
    [pem_pos=1, final_scores=0.6258163,
final_local_scores=0.5806201, pem=-2.103781,
similarity=32.85726]
21: 248083 Norwich_City_F.C.
    [pem_pos=2, final_scores=0.7842739,
final_local_scores=0.74585336, pem=-2.7278075,
similarity=41.181705]
22: 217730 Wolverhampton_Wanderers_F.C.
    [pem_pos=1, final_scores=0.90041083,
final_local_scores=0.87182885, pem=-2.274954,
similarity=42.47838]
23: 451131 Oldham_Athletic_A.F.C.
    [pem_pos=3, final_scores=0.6820733,
final_local_scores=0.6355944, pem=-2.8632221,
similarity=38.627766]
24: 352843 Ipswich_Town_F.C.
    [pem_pos=4, final_scores=0.8947056,
final_local_scores=0.87042266, pem=-2.6931517,
similarity=44.71064]
25: 203456 Port_Vale_F.C.
    [pem_pos=0, final_scores=0.626983,
final_local_scores=0.5853521, pem=0.0, similarity=21.557787]
26: 451165 Oxford_United_F.C.
    [pem_pos=18, final_scores=0.8777376,
final_local_scores=0.85409987, pem=-4.3872933,
similarity=53.43654]
27: 168007 Reading_F.C.
    [pem_pos=0, final_scores=1.0493267,
final_local_scores=1.0357628, pem=-1.2007197,
similarity=41.52864]
28: 203434 Stoke_City_F.C.
    [pem_pos=0, final_scores=0.58846897,
final_local_scores=0.5361029, pem=-0.3417574,
similarity=21.94662]
```

29: 451168 Southend_United_F.C.
[pem_pos=2, final_scores=0.71434444,
final_local_scores=0.6710871, pem=-1.771178,
similarity=33.74823]
30: 74291 Swindon_Town_F.C.
[pem_pos=4, final_scores=0.62064856,
final_local_scores=0.572105, pem=-2.489671,
similarity=34.701687]
31: 9316 England
[pem_pos=1, final_scores=0.6930856,
final_local_scores=0.75767523, pem=-1.3527626,
similarity=34.056786]
32: 245357 Blackpool_F.C.
[pem_pos=1, final_scores=0.85150427,
final_local_scores=0.82083744, pem=-0.8537509,
similarity=33.22797]
33: 451136 Wycombe_Wanderers_F.C.
[pem_pos=1, final_scores=0.6829106,
final_local_scores=0.6482634, pem=-1.147249,
similarity=29.674105]
34: 451121 A.F.C._Bournemouth
[pem_pos=1, final_scores=0.7221201,
final_local_scores=0.6797914, pem=-1.0576377,
similarity=30.127642]
35: 451132 Peterborough_United_F.C.
[pem_pos=6, final_scores=0.85560125,
final_local_scores=0.829013, pem=-3.1818013,
similarity=46.132244]
36: 241161 Bristol_Rovers_F.C.
[pem_pos=0, final_scores=0.46422392,
final_local_scores=0.39672548, pem=0.0, similarity=15.928793]
37: 451135 Stockport_County_F.C.
[pem_pos=4, final_scores=0.6166668,
final_local_scores=0.5706286, pem=-2.5287716,
similarity=34.870262]
38: 451156 Bury_F.C.
[pem_pos=0, final_scores=0.78014416,
final_local_scores=0.7402303, pem=-0.72067994,
similarity=30.098831]
39: 4392 Bristol_City_F.C.
[pem_pos=0, final_scores=0.47278923,
final_local_scores=0.4076376, pem=0.0, similarity=16.254433]
40: 69438 Crewe_Alexandra_F.C.
[pem_pos=3, final_scores=0.5983709,
final_local_scores=0.5497606, pem=-2.8374243,
similarity=35.926018]
41: 85099 Watford_F.C.
[pem_pos=1, final_scores=0.56778306,
final_local_scores=0.51137346, pem=-0.95984375,
similarity=24.569897]
42: 165278 Gillingham_F.C.
[pem_pos=0, final_scores=0.5935014,

```
final_local_scores=0.54184264, pem=-0.2568813,
similarity=21.656338]
43: 451126 Chesterfield_F.C.
    [pem_pos=1, final_scores=0.76239914,
final_local_scores=0.7257942, pem=-1.2678006,
similarity=32.643356]
44: 18367 Luton_Town_F.C.
    [pem_pos=2, final_scores=0.6101688,
final_local_scores=0.5558, pem=-2.5537767,
similarity=34.56373]
45: 386362 Rotherham_United_F.C.
    [pem_pos=3, final_scores=0.5980581,
final_local_scores=0.5457939, pem=-2.6427798,
similarity=34.749138]
46: 435024 Millwall_F.C.
    [pem_pos=0, final_scores=0.5862233,
final_local_scores=0.53046155, pem=0.0, similarity=19.919744]
47: 376725 Burnley_F.C.
    [pem_pos=1, final_scores=0.37212437,
final_local_scores=0.29692435, pem=-0.9914584,
similarity=18.342234]
48: 448599 Notts_County_F.C.
    [pem_pos=0, final_scores=0.58691734,
final_local_scores=0.53110224, pem=0.0, similarity=19.938864]
49: 429092 Preston_North_End_F.C.
    [pem_pos=4, final_scores=0.9513338,
final_local_scores=0.9246023, pem=-3.194944,
similarity=49.056293]
50: 451207 Shrewsbury_Town_F.C.
    [pem_pos=12, final_scores=0.9760173,
final_local_scores=0.96057004, pem=-3.2502756,
similarity=50.430542]
51: 451122 Brentford_F.C.
    [pem_pos=0, final_scores=0.8137767,
final_local_scores=0.7827245, pem=-0.36028567,
similarity=29.407063]
52: 359501 Walsall_F.C.
    [pem_pos=1, final_scores=0.65453786,
final_local_scores=0.6051033, pem=-0.61973774,
similarity=25.517433]
53: 413286 Wrexham_F.C.
    [pem_pos=1, final_scores=0.60021925,
final_local_scores=0.5517558, pem=-0.88259363,
similarity=25.354887]
54: 9316 England
    [pem_pos=1, final_scores=0.6198655,
final_local_scores=0.6776934, pem=-1.3527626,
similarity=31.669968]
55: 69455 Brighton_&Hove_Albion_F.C.
    [pem_pos=3, final_scores=1.0343623,
final_local_scores=1.0235493, pem=-3.1271749,
similarity=51.64053]
```

56: 451167 Scunthorpe_United_F.C.
[pem_pos=1, final_scores=0.7250233,
final_local_scores=0.6854549, pem=-1.507785,
similarity=32.74462]
57: 5624949 Cardiff_City_F.C.
[pem_pos=3, final_scores=0.8579008,
final_local_scores=0.8272583, pem=-3.4864018,
similarity=47.736347]
58: 451127 Colchester_United_F.C.
[pem_pos=7, final_scores=0.8863707,
final_local_scores=0.8735631, pem=-2.8568413,
similarity=45.694527]
59: 451201 Hereford_United_F.C.
[pem_pos=6, final_scores=1.1801195,
final_local_scores=1.1911097, pem=-2.7578042,
similarity=54.63217]
60: 451160 Doncaster_Rovers_F.C.
[pem_pos=6, final_scores=0.6272221,
final_local_scores=0.576678, pem=-2.950685,
similarity=37.34522]
61: 451159 Darlington_F.C.
[pem_pos=1, final_scores=0.6772445,
final_local_scores=0.631435, pem=-1.8695662,
similarity=33.099983]
62: 11228 Fulham_F.C.
[pem_pos=0, final_scores=0.54882497,
final_local_scores=0.49889594, pem=-0.3209339,
similarity=20.723051]
63: 451157 Carlisle_United_F.C.
[pem_pos=12, final_scores=0.93395144,
final_local_scores=0.90939623, pem=-3.5999887,
similarity=50.80521]
64: 451358 Hull_City_A.F.C.
[pem_pos=11, final_scores=0.7261389,
final_local_scores=0.6859881, pem=-3.7682815,
similarity=45.053467]
65: 451173 Barnet_F.C.
[pem_pos=0, final_scores=0.74307865,
final_local_scores=0.71982455, pem=-0.6797195,
similarity=29.267134]
66: 451162 Leyton_Orient_F.C.
[pem_pos=0, final_scores=0.47761875,
final_local_scores=0.41468322, pem=0.0, similarity=16.464687]
67: 451130 Hartlepool_United_F.C.
[pem_pos=1, final_scores=0.7808195,
final_local_scores=0.7426345, pem=-1.75166,
similarity=35.777206]
68: 451164 Mansfield_Town_F.C.
[pem_pos=16, final_scores=0.84113926,
final_local_scores=0.8148611, pem=-4.267106,
similarity=51.61198]
69: 451166 Rochdale_A.F.C.

[pem_pos=1, final_scores=0.97578067,
final_local_scores=0.9581111, pem=-0.95097107,
similarity=37.853188]
70: 946449 Scarborough_F.C.
[pem_pos=2, final_scores=0.8624962,
final_local_scores=0.8499157, pem=-1.8160111,
similarity=39.328644]
71: 425750 Northampton_Town_F.C.
[pem_pos=10, final_scores=0.93691677,
final_local_scores=0.9169155, pem=-3.2381344,
similarity=49.06178]
72: 193049 Torquay_United_F.C.
[pem_pos=2, final_scores=0.7268783,
final_local_scores=0.6870367, pem=-2.019306,
similarity=35.57356]
73: 367030 Wigan_Athletic_F.C.
[pem_pos=2, final_scores=0.5390186,
final_local_scores=0.4858815, pem=-2.32737,
similarity=31.245985]
74: 451195 Chester_City_F.C.
[pem_pos=1, final_scores=0.7090793,
final_local_scores=0.68010825, pem=-3.2160804,
similarity=41.875046]
75: 26994 Scotland
[pem_pos=0, final_scores=0.46287626,
final_local_scores=0.5108146, pem=-0.47951394,
similarity=21.94111]
76: 451245 East_Fife_F.C.
[pem_pos=0, final_scores=0.7468781,
final_local_scores=0.7352419, pem=0.0, similarity=26.030798]
77: 1347678 Clydebank_F.C.
[pem_pos=1, final_scores=0.48637688,
final_local_scores=0.4557019, pem=-1.3275584,
similarity=24.908236]
78: 367890 Greenock_Morton_F.C.
[pem_pos=0, final_scores=0.5887946,
final_local_scores=0.5583944, pem=0.0, similarity=20.753315]
79: 226415 Partick_Thistle_F.C.
[pem_pos=2, final_scores=0.6372948,
final_local_scores=0.6055981, pem=-1.2851555,
similarity=29.150845]
80: 783089 St._Mirren_F.C.
[pem_pos=0, final_scores=0.6660548,
final_local_scores=0.63828444, pem=0.0, similarity=23.137394]
81: 393756 St._Johnstone_F.C.
[pem_pos=0, final_scores=0.82056576,
final_local_scores=0.8037346, pem=0.0, similarity=28.074757]
82: 1216952 Airdrieonians_F.C._(1878)
[pem_pos=1, final_scores=0.24408673,
final_local_scores=0.18483946, pem=-1.8786064,
similarity=19.821844]
83: 306504 Dundee_F.C.

[pem_pos=1, final_scores=0.651212,
final_local_scores=0.61693114, pem=-1.3661808,
similarity=29.929672]
84: 26994 Scotland
[pem_pos=0, final_scores=0.45421094,
final_local_scores=0.5013491, pem=-0.47951394,
similarity=21.658642]
85: 451240 Ayr_United_F.C.
[pem_pos=17, final_scores=0.58911985,
final_local_scores=0.5520886, pem=-3.340128,
similarity=38.729275]
86: 451243 Clyde_F.C.
[pem_pos=0, final_scores=0.9211764,
final_local_scores=0.9140118, pem=-0.9832426,
similarity=36.712673]
87: 880354 Queen_of_the_South_F.C.
[pem_pos=0, final_scores=0.73676366,
final_local_scores=0.72053695, pem=0.0, similarity=25.591972]
88: 443155 Dumbarton_F.C.
[pem_pos=1, final_scores=0.6655138,
final_local_scores=0.65051347, pem=-1.254432,
similarity=30.324127]
89: 451241 Brechin_City_F.C.
[pem_pos=2, final_scores=0.7038534,
final_local_scores=0.6998851, pem=-1.9178002,
similarity=35.404976]
90: 451183 Livingston_F.C.
[pem_pos=0, final_scores=0.97456986,
final_local_scores=0.98696, pem=-1.8614511,
similarity=43.66543]
91: 653922 Stenhousemuir_F.C.
[pem_pos=0, final_scores=0.6705332,
final_local_scores=0.658842, pem=0.0, similarity=23.750874]
92: 880391 Stranraer_F.C.
[pem_pos=1, final_scores=0.5614199,
final_local_scores=0.540427, pem=-1.0295545,
similarity=25.816011]
93: 26994 Scotland
[pem_pos=0, final_scores=0.47914898,
final_local_scores=0.5285901, pem=-0.47951394,
similarity=22.471567]
94: 451244 Cowdenbeath_F.C.
[pem_pos=0, final_scores=0.5993865,
final_local_scores=0.57929146, pem=-0.24741817,
similarity=22.722424]
95: 451238 Arbroath_F.C.
[pem_pos=1, final_scores=0.88300735,
final_local_scores=0.882997, pem=-0.8612016,
similarity=35.12345]
96: 817230 East_Stirlingshire_F.C.
[pem_pos=0, final_scores=0.8293409,
final_local_scores=0.83730817, pem=0.0, similarity=29.076658]

```

97: 743540 Inverness_Caledonian_Thistle_F.C.
    [pem_pos=6, final_scores=0.92265636,
final_local_scores=0.9257917, pem=-3.2290676,
similarity=49.277355]
98: 451237 Alloa_Athletic_F.C.
    [pem_pos=1, final_scores=1.0306737,
final_local_scores=1.0492191, pem=-0.79396856,
similarity=39.71823]
99: 880408 Montrose_F.C.
    [pem_pos=1, final_scores=0.71569836,
final_local_scores=0.7144006, pem=-2.1839786,
similarity=37.285667]
100: 880380 Ross_County_F.C.
    [pem_pos=0, final_scores=0.684838,
final_local_scores=0.66816646, pem=0.0, similarity=24.029133]
101: 408985 Forfar_Athletic_F.C.
    [pem_pos=1, final_scores=0.7276744,
final_local_scores=0.721003, pem=-1.7646968,
similarity=35.202576]
102: 9316 England
    [pem_pos=1, final_scores=0.5605708,
final_local_scores=0.61292297, pem=-1.3527626,
similarity=29.73709]
103: 84588 Queens_Park_Rangers_F.C.
    [pem_pos=0, final_scores=0.5851731,
final_local_scores=0.5340056, pem=0.0, similarity=20.025505]
104: 298601 Bolton_Wanderers_F.C.
    [pem_pos=2, final_scores=0.6620521,
final_local_scores=0.61255044, pem=-2.2845325,
similarity=34.793083]

```

Sample: b'417_SOCCER'

SOCCER - [1 1 COSTA RICA 1] AND [2 2 CHILE 2] DRAW 1-1 IN FRIENDLY . [3 3 LIBERIA 3] , [4 4 Costa Rica 4] 1996-08-26 [5 5 Costa Rica 5] and [6 6 Chile 6] drew 1-1 (halftime 1-0) in a friendly soccer international on Sunday . Scorers : [7 7 Costa Rica 7] - Ronaldo Gonzalez (10th minute , penalty [8 8 Chile 8] - [9 9 Marcelo Salas 9] (80th) Attendance : 8,000

```

1: 1014478 Costa_Rica_national_football_team
2: 697710 Chile_national_football_team
3: 17791 Liberia
4: 5551 Costa_Rica
5: 1014478 Costa_Rica_national_football_team
6: 697710 Chile_national_football_team
7: 1014478 Costa_Rica_national_football_team
8: 697710 Chile_national_football_team
9: 424153 Marcelo_Salas
1: 1014478 Costa_Rica_national_football_team
    [pem_pos=1, final_scores=0.62937516,
final_local_scores=0.5891634, pem=-1.9901731,
similarity=32.494396]
2: 697710 Chile_national_football_team

```

```
[pem_pos=1, final_scores=0.63712627,
final_local_scores=0.5954928, pem=-2.4683056,
similarity=35.283436]
3: 17791 Liberia
    [pem_pos=0, final_scores=0.37614352,
final_local_scores=0.46442747, pem=-0.26536247,
similarity=19.392237]
4: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.4609924,
final_local_scores=0.5427296, pem=-0.30652362,
similarity=21.95277]
5: 1014478 Costa_Rica_national_football_team
    [pem_pos=1, final_scores=0.52122086,
final_local_scores=0.4710213, pem=-1.9901731,
similarity=28.9688]
6: 697710 Chile_national_football_team
    [pem_pos=1, final_scores=0.6544214,
final_local_scores=0.61438507, pem=-2.4683056,
similarity=35.847218]
7: 1014478 Costa_Rica_national_football_team
    [pem_pos=1, final_scores=0.5809496,
final_local_scores=0.53626585, pem=-1.9901731,
similarity=30.915825]
8: 697710 Chile_national_football_team
    [pem_pos=1, final_scores=0.6548287,
final_local_scores=0.61482996, pem=-2.4683056,
similarity=35.860493]
9: 424153 Marcelo_Salas
    [pem_pos=0, final_scores=0.66189754,
final_local_scores=0.7191512, pem=0.0, similarity=25.55062]
```

Sample: b'418_SOCCER'

SOCCER - [1 1 SEYCHELLES 1] FAIL IN BID FOR HISTORIC
VICTORY . [2 2 Mark Gleeson 2] [3 3 JOHANNESBURG 3]
1996-08-26 The tiny islands of the [4 4 Seychelles 4] failed
to make soccer history at the weekend when they bowed out of
the preliminary rounds of the [5 5 African Nations Cup 5] .
Trailing fellow Indian Ocean islanders [6 6 Mauritius 6] 1-0
from the first leg , they were held to a 1-1 draw at home on
Saturday despite playing against 10 men for most of the
match . The 2-1 aggregate took [7 7 Mauritius 7] into the
group phase of the qualifiers for the 1998 finals , and kept
up the [8 8 Seychelles 8] ' record of never having won an
official match in their 10 years of [9 9 FIFA 9] membership .
The [10 10 Seychelles 10] must have thought they were on
course for a historic breakthrough when [11 11 Mauritian 11]
midfielder Andre Caboche was sent off for a crude tackle in
the 19th minute . But the visitors responded to the setback
immediately -- veteran striker Ashley Mocude scoring a minute
later to give them a two-goal aggregate lead . Although Danny
Rose 's 50th-minute equaliser gave the Seychellois renewed
hope they could not find the net again and were eliminated . [

[12 12 Mauritius 12] now play in group seven of the qualifiers against [13 13 Malawi 13] , [14 14 Mozambique 14] and favourites [15 15 Zambia 15] . [16 16 Namibia 16] , who drew 0-0 with [17 17 Botswana 17] in their first leg , won the second leg in [18 18 Windhoek 18] 6-0 to stretch their unbeaten run to eight matches and continue their remarkable progress on the [19 19 African 19] soccer stage . They now play in group five with [20 20 Cameroon 20] , [21 21 Gabon 21] and [22 22 Kenya 22] . German-based striker [23 23 Bachirou Salou 23] returned home to [24 24 Togo 24] to score the decisive only goal of their tie against [25 25 Congo 25] . [26 26 Salou 26] , who plays for [27 27 MSV Duisburg 27] in the [28 28 Bundesliga 28] , scored in the 53rd minute of Sunday 's match in [29 29 Lome 29] for a 1-0 aggregate win which takes his side into group six , where they will meet [30 30 Liberia 30] , [31 31 Tanzania 31] and [32 32 Zaire 32] . [33 33 Ethiopia 33] needed a penalty shoot-out in Addis Ababa to overcome [34 34 Uganda 34] after a 2-2 aggregate scoreline . Both legs ended 1-1 before [35 35 Ethiopia 35] won the spot kick decider 4-2 . [36 36 Uganda 36] 's elimination follows their humiliating 5-1 aggregate defeat by [37 37 Angola 37] in June 's [38 38 World Cup 38] qualifying preliminaries . The other preliminary round second leg match , between [39 39 Mauritania 39] and [40 40 Benin 40] in [41 41 Nouakchott 41] , was postponed until Friday . [42 42 Benin 42] won the first leg 4-1 .

1: 27288 Seychelles
2: 25873962 Mark_Gleeson_(journalist)
3: 58893 Johannesburg
4: 27288 Seychelles
5: 250680 Africa_Cup_of_Nations
6: 1205110 Mauritius_national_football_team
7: 1205110 Mauritius_national_football_team
8: 1263594 Seychelles_national_football_team
9: 11049 FIFA
10: 1263594 Seychelles_national_football_team
11: 19201 Mauritius
12: 1205110 Mauritius_national_football_team
13: 1198154 Malawi_national_football_team
14: 1132599 Mozambique_national_football_team
15: 1045380 Zambia_national_football_team
16: 1199910 Namibia_national_football_team
17: 1158358 Botswana_national_football_team
18: 57669 Windhoek
19: 5334607 Africa
20: 1005798 Cameroon_national_football_team
21: 1109291 Gabon_national_football_team
22: 1105751 Kenya_national_football_team
23: 14716002 Bachirou_Salou
24: 30138 Togo
25: 1047784 Congo_national_football_team
26: 14716002 Bachirou_Salou

```
27: 1586579 MSV_Duisburg
28: 686142 Bundesliga
29: 57067 Lomé
30: 1080283 Liberia_national_football_team
31: 1202658 Tanzania_national_football_team
32: 1030243 DR_Congo_national_football_team
33: 1089708 Ethiopia_national_football_team
34: 1092055 Uganda_national_football_team
35: 1089708 Ethiopia_national_football_team
36: 1092055 Uganda_national_football_team
37: 1083920 Angola_national_football_team
38: 126460 1998_FIFA_World_Cup
39: 1234981 Mauritania_national_football_team
40: 1198472 Benin_national_football_team
41: 51592 Nouakchott
42: 1198472 Benin_national_football_team
1: 27288 Seychelles
    [pem_pos=0, final_scores=0.35937172,
final_local_scores=0.4192956, pem=-0.3716595,
similarity=18.623472]
2: 25873962 Mark_Gleeson_(journalist)
    [pem_pos=1, final_scores=0.30967754,
final_local_scores=0.46681422, pem=-0.30748472,
similarity=19.69253]
3: 58893 Johannesburg
    [pem_pos=0, final_scores=0.44751674,
final_local_scores=0.4925077, pem=-0.09562059,
similarity=19.307125]
4: 27288 Seychelles
    [pem_pos=0, final_scores=0.40682137,
final_local_scores=0.47112715, pem=-0.3716595,
similarity=20.170229]
5: 250680 Africa_Cup_of_Nations
    [pem_pos=0, final_scores=0.54452944,
final_local_scores=0.51909906, pem=0.0, similarity=19.580664]
6: 1205110 Mauritius_national_football_team
    [pem_pos=1, final_scores=0.58815306,
final_local_scores=0.5902818, pem=-2.5433543,
similarity=35.536057]
7: 1205110 Mauritius_national_football_team
    [pem_pos=1, final_scores=0.5426728,
final_local_scores=0.54060143, pem=-2.5433543,
similarity=34.053493]
8: 1263594 Seychelles_national_football_team
    [pem_pos=1, final_scores=0.53750867,
final_local_scores=0.5454894, pem=-2.1369176,
similarity=31.989094]
9: 11049 FIFA
    [pem_pos=0, final_scores=0.31339914,
final_local_scores=0.29993805, pem=0.0, similarity=13.0404625]
10: 1263594 Seychelles_national_football_team
    [pem_pos=1, final_scores=0.47453898,
```

```
final_local_scores=0.47670454, pem=-2.1369176,
similarity=29.936417]
11: 19201 Mauritius
    [pem_pos=0, final_scores=0.3621034,
final_local_scores=0.40891403, pem=-0.54702646,
similarity=19.26734]
12: 1205110 Mauritius_national_football_team
    [pem_pos=1, final_scores=0.678278,
final_local_scores=0.6887296, pem=-2.5433543,
similarity=38.473934]
13: 1198154 Malawi_national_football_team
    [pem_pos=1, final_scores=0.7405475,
final_local_scores=0.75597703, pem=-2.1824102,
similarity=38.51786]
14: 1132599 Mozambique_national_football_team
    [pem_pos=1, final_scores=0.48602337,
final_local_scores=0.48261732, pem=-2.788136,
similarity=33.654293]
15: 1045380 Zambia_national_football_team
    [pem_pos=1, final_scores=0.94137555,
final_local_scores=0.9612531, pem=-2.1802778,
similarity=44.632114]
16: 1199910 Namibia_national_football_team
    [pem_pos=2, final_scores=0.5508387,
final_local_scores=0.5463987, pem=-2.6771367,
similarity=34.954025]
17: 1158358 Botswana_national_football_team
    [pem_pos=1, final_scores=0.5941562,
final_local_scores=0.60047513, pem=-2.4778914,
similarity=35.48425]
18: 57669 Windhoek
    [pem_pos=0, final_scores=0.42369282,
final_local_scores=0.48968428, pem=0.0, similarity=18.70287]
19: 5334607 Africa
    [pem_pos=0, final_scores=0.44120145,
final_local_scores=0.50879556, pem=-0.91165566,
similarity=24.230913]
20: 1005798 Cameroon_national_football_team
    [pem_pos=1, final_scores=0.80915755,
final_local_scores=0.8083093, pem=-1.8415786,
similarity=38.226067]
21: 1109291 Gabon_national_football_team
    [pem_pos=1, final_scores=0.8164578,
final_local_scores=0.82642794, pem=-2.1652567,
similarity=40.526974]
22: 1105751 Kenya_national_football_team
    [pem_pos=2, final_scores=0.41426736,
final_local_scores=0.39338815, pem=-3.2429898,
similarity=33.465084]
23: 14716002 Bachirou_Salou
    [pem_pos=0, final_scores=0.664737,
final_local_scores=0.7322159, pem=0.0, similarity=25.940495]
```

```
24: 30138 Togo
    [pem_pos=0, final_scores=0.43825585,
final_local_scores=0.4797275, pem=-0.4966172,
similarity=21.106419]
25: 1047784 Congo_national_football_team
    [pem_pos=2, final_scores=0.789776,
final_local_scores=0.81255215, pem=-2.414126,
similarity=41.46628]
26: 14716002 Bachirou_Salou
    [pem_pos=1, final_scores=0.6433614,
final_local_scores=0.71742934, pem=0.0, similarity=25.499235]
27: 1586579 MSV_Duisburg
    [pem_pos=0, final_scores=0.55415434,
final_local_scores=0.6009791, pem=0.0, similarity=22.024128]
28: 686142 Bundesliga
    [pem_pos=0, final_scores=0.36127096,
final_local_scores=0.37336767, pem=-0.08715299,
similarity=15.7057]
29: 57067 Lomé
    [pem_pos=0, final_scores=0.3702727,
final_local_scores=0.39483964, pem=0.0, similarity=15.872515]
30: 1080283 Liberia_national_football_team
    [pem_pos=1, final_scores=0.6696132,
final_local_scores=0.6910716, pem=-2.2176507,
similarity=36.7726]
31: 1202658 Tanzania_national_football_team
    [pem_pos=1, final_scores=0.7895877,
final_local_scores=0.82069844, pem=-2.2414286,
similarity=40.77023]
32: 1030243 DR_Congo_national_football_team
    [pem_pos=1, final_scores=0.8907812,
final_local_scores=0.8928656, pem=-1.6428272,
similarity=39.668552]
33: 1089708 Ethiopia_national_football_team
    [pem_pos=1, final_scores=0.68906313,
final_local_scores=0.68447804, pem=-2.7670166,
similarity=39.56337]
34: 1092055 Uganda_national_football_team
    [pem_pos=1, final_scores=0.7052349,
final_local_scores=0.7091718, pem=-2.80235,
similarity=40.492428]
35: 1089708 Ethiopia_national_football_team
    [pem_pos=1, final_scores=0.7419705,
final_local_scores=0.7422713, pem=-2.7670166,
similarity=41.288036]
36: 1092055 Uganda_national_football_team
    [pem_pos=1, final_scores=0.6032593,
final_local_scores=0.5977789, pem=-2.80235,
similarity=37.168243]
37: 1083920 Angola_national_football_team
    [pem_pos=1, final_scores=0.6551133,
final_local_scores=0.6488666, pem=-2.6096098,
```

```

similarity=37.644653]
38: 126460 1998_FIFA_World_Cup
    [pem_pos=18, final_scores=0.2939198,
final_local_scores=0.26481068, pem=-4.2172847,
similarity=34.92644]
39: 1234981 Mauritania_national_football_team
    [pem_pos=1, final_scores=0.5513902,
final_local_scores=0.56787044, pem=-2.1465292,
similarity=32.70926]
40: 1198472 Benin_national_football_team
    [pem_pos=1, final_scores=0.66071296,
final_local_scores=0.67061824, pem=-2.165171,
similarity=35.876835]
41: 51592 Nouakchott
    [pem_pos=0, final_scores=0.3079474,
final_local_scores=0.38186967, pem=0.0, similarity=15.485466]
42: 1198472 Benin_national_football_team
    [pem_pos=1, final_scores=0.6274837,
final_local_scores=0.6343203, pem=-2.165171,
similarity=34.793633]

```

Sample: b'419_SOCCER'

SOCCER - [1 1 AFRICAN NATIONS CUP 1] COLLATED RESULTS . [2 2 JOHANNESBURG 2] 1996-08-26 Collated results of [3 3 African Nations Cup 3] preliminary round , second leg matches played at the weekend : [4 4 Ethiopia 4] 1 [5 5 Uganda 5] 1 2-2 on aggregate . [6 6 Ethiopia 6] win 4-2 on penalties [7 7 Mauritania 7] v [8 8 Benin 8] postponed to Friday [9 9 Benin 9] lead 4-1 from the first leg [10 10 Namibia 10] 6 [11 11 Botswana 11] 0 [12 12 Namibia 12] win 6-0 on aggregate [13 13 Seychelles 13] 1 [14 14 Mauritius 14] 1 [15 15 Mauritius 15] win 2-1 on aggregate [16 16 Togo 16] 1 [17 17 Congo 17] 0 [18 18 Togo 18] win 1-0 on aggregate [19 19 Central African Republic 19] walkover v [20 20 Burundi 20] Winners progress to qualifying groups to start in October .

```

1: 250680 Africa_Cup_of_Nations
2: 58893 Johannesburg
3: 250680 Africa_Cup_of_Nations
4: 1089708 Ethiopia_national_football_team
5: 1092055 Uganda_national_football_team
6: 1089708 Ethiopia_national_football_team
7: 1234981 Mauritania_national_football_team
8: 1198472 Benin_national_football_team
9: 1198472 Benin_national_football_team
10: 1199910 Namibia_national_football_team
11: 1158358 Botswana_national_football_team
12: 1199910 Namibia_national_football_team
13: 1263594 Seychelles_national_football_team
14: 1205110 Mauritius_national_football_team
15: 1205110 Mauritius_national_football_team
16: 1080720 Togo_national_football_team
17: 1047784 Congo_national_football_team

```

```
18: 1080720 Togo_national_football_team
19: 1263067 Central_African_Republic_national_football_team
20: 1212797 Burundi_national_football_team
1: 250680 Africa_Cup_of_Nations
    [pem_pos=0, final_scores=0.6573076,
final_local_scores=0.63774115, pem=0.0, similarity=23.121181]
2: 58893 Johannesburg
    [pem_pos=0, final_scores=0.44903517,
final_local_scores=0.5228457, pem=-0.09562059,
similarity=20.212471]
3: 250680 Africa_Cup_of_Nations
    [pem_pos=0, final_scores=0.52424175,
final_local_scores=0.49238688, pem=0.0, similarity=18.78352]
4: 1089708 Ethiopia_national_football_team
    [pem_pos=1, final_scores=0.7120157,
final_local_scores=0.7045268, pem=-2.7670166,
similarity=40.161663]
5: 1092055 Uganda_national_football_team
    [pem_pos=1, final_scores=0.7098014,
final_local_scores=0.709527, pem=-2.80235,
similarity=40.50303]
6: 1089708 Ethiopia_national_football_team
    [pem_pos=1, final_scores=0.70789546,
final_local_scores=0.70002604, pem=-2.7670166,
similarity=40.02735]
7: 1234981 Mauritania_national_football_team
    [pem_pos=1, final_scores=0.44902784,
final_local_scores=0.4492255, pem=-2.1465292,
similarity=29.168655]
8: 1198472 Benin_national_football_team
    [pem_pos=1, final_scores=0.8416737,
final_local_scores=0.8541258, pem=-2.165171,
similarity=41.353065]
9: 1198472 Benin_national_football_team
    [pem_pos=1, final_scores=0.72005916,
final_local_scores=0.7212804, pem=-2.165171,
similarity=37.388695]
10: 1199910 Namibia_national_football_team
    [pem_pos=2, final_scores=0.61392313,
final_local_scores=0.6053062, pem=-2.6771367,
similarity=36.711945]
11: 1158358 Botswana_national_football_team
    [pem_pos=1, final_scores=0.54819876,
final_local_scores=0.54524964, pem=-2.4778914,
similarity=33.83621]
12: 1199910 Namibia_national_football_team
    [pem_pos=2, final_scores=0.53557616,
final_local_scores=0.5197241, pem=-2.6771367,
similarity=34.158]
13: 1263594 Seychelles_national_football_team
    [pem_pos=1, final_scores=0.51272583,
final_local_scores=0.51679903, pem=-2.1369176,
```

```

similarity=31.132917]
14: 1205110 Mauritius_national_football_team
    [pem_pos=1, final_scores=0.61934143,
final_local_scores=0.62595123, pem=-2.5433543,
similarity=36.600502]
15: 1205110 Mauritius_national_football_team
    [pem_pos=1, final_scores=0.63713783,
final_local_scores=0.64539105, pem=-2.5433543,
similarity=37.180626]
16: 1080720 Togo_national_football_team
    [pem_pos=1, final_scores=0.5597158,
final_local_scores=0.5482872, pem=-2.2790594,
similarity=32.845577]
17: 1047784 Congo_national_football_team
    [pem_pos=2, final_scores=0.9046858,
final_local_scores=0.93130547, pem=-2.414126,
similarity=45.010117]
18: 1080720 Togo_national_football_team
    [pem_pos=1, final_scores=0.5533151,
final_local_scores=0.54129547, pem=-2.2790594,
similarity=32.63693]
19: 1263067 Central_African_Republic_national_football_team
    [pem_pos=1, final_scores=0.4175496,
final_local_scores=0.4176131, pem=-2.0891073,
similarity=27.91301]
20: 1212797 Burundi_national_football_team
    [pem_pos=1, final_scores=0.61106855,
final_local_scores=0.63923746, pem=-2.2049255,
similarity=35.156563]

```

Sample: b'420_SOCCER'

```

SOCCER - [1 1 UKRAINIAN 1] PREMIER DIVISION RESULTS /
STANDINGS . [2 2 KIEV 2] 1996-08-26 Results of [3 3 Ukraine 3]
premier division matches played at the weekend : [4 4 Dynamo K
iev 4] 5 [5 5 Kremin Kremenchuk 5] 0 [6 6 Vorskla Poltava 6]
2 [7 7 Nyva Ternopil 7] 1 [8 8 Torpedo Zaporizhya 8] 2 [9 9 Sh
akhtar Donetsk 9] 1 Kryvbas Kryvy Rig 1 [10 10 Karpaty Lviv 10]
] 2 Prykarpatty Ivano-Frankivsk 0 Zirka-Nibas Kirovohrad 0 [1
1 11 Chornomorets Odessa 11] 2 [12 12 Metalurg Zaporizhya 12]
1 [13 13 Dnipro Dnipropetrovsk 13] 2 [14 14 CSKA Kiev 14] 1
Standings ( tabulated under played , won , drawn , lost ,
goals for , against , points ) : [15 15 Dynamo 15] 6 5 0 1 16
2 15 [16 16 Vorskla 16] 6 4 2 0 11 3 14 [17 17 Dnipro 17] 6 4
1 1 13 6 13 [18 18 Chornomorets 18] 6 4 1 1 11 7 13 [19 19 Sh
akhtar 19] 6 3 2 1 10 3 11 [20 20 Metalurg 20] 6 3 2 1 9 6 11
[21 21 Karpaty 21] 6 3 1 2 9 5 10 Zirka-Nibas 6 3 1 2 6 8 10 [
22 22 Torpedo 22] 6 3 1 2 8 7 10 [23 23 Tavria 23] 5 2 0 3 3
7 6 [24 24 Nyva Ternopil 24] 6 2 0 4 4 11 6 1 CSKA 6 1 1 4 4
7 4 [25 25 Kryvbas 25] 6 1 1 4 5 9 4 Nyva Vinnytsya 5 0 2 3 1
7 2 Prykarpatty 6 0 2 4 4 13 2 [26 26 Kremin 26] 6 0 1 5 1
14 1
1: 31750 Ukraine

```

2: 585629 Kiev
3: 31750 Ukraine
4: 889025 FC_Dynamo_Kyiv
5: 7009763 MFC_Kremin_Kremenchuk
6: 3536552 FC_Vorskla_Poltava
7: 7010381 FC_Nyva_Ternopil
8: 7011579 FC_Torpedo_Zaporizhya
9: 861305 FC_Shakhtar_Donetsk
10: 4388256 FC_Karpaty_Lviv
11: 3435061 FC_Chornomorets_Odesa
12: 3517746 FC_Metalurh_Zaporizhya
13: 895551 FC_Dnipro_Dnipropetrovsk
14: 5056113 FC_CSKA_Kyiv
15: 889025 FC_Dynamo_Kyiv
16: 3536552 FC_Vorskla_Poltava
17: 895551 FC_Dnipro_Dnipropetrovsk
18: 3435061 FC_Chornomorets_Odesa
19: 861305 FC_Shakhtar_Donetsk
20: 17864564 FK_Metalurg_Skopje
21: 4388256 FC_Karpaty_Lviv
22: 1443555 FC_Torpedo_Moscow
23: 3325822 SC_Tavriya_Simferopol
24: 7010381 FC_Nyva_Ternopil
25: 3800152 FC_Kryvbas_Kryvyi_Rih
26: 7009763 MFC_Kremin_Kremenchuk

1: 5056113 FC_CSKA_Kyiv
[winner: 5056113 FC_CSKA_Kyiv, pem_pos=7,
final_scores=0.013671204, final_local_scores=-0.012916267,
pem=-2.812644, similarity=18.999863]
gt: 5056113 FC_CSKA_Kyiv, pem_pos=7,
final_scores=0.013671204, final_local_scores=-0.012916267,
pem=-2.812644, similarity=18.999863]

1: 31750 Ukraine
[pem_pos=1, final_scores=0.62807196,
final_local_scores=0.67673725, pem=-1.0458583,
similarity=29.972443]

2: 585629 Kiev
[pem_pos=0, final_scores=0.4524247,
final_local_scores=0.48334408, pem=-0.032242518,
similarity=18.689003]

3: 31750 Ukraine
[pem_pos=0, final_scores=0.5622666,
final_local_scores=0.6048549, pem=-0.35407752,
similarity=24.065317]

4: 889025 FC_Dynamo_Kyiv
[pem_pos=0, final_scores=0.6016852,
final_local_scores=0.5603471, pem=0.0, similarity=20.811588]

5: 7009763 MFC_Kremin_Kremenchuk
[pem_pos=0, final_scores=0.5588544,
final_local_scores=0.5691603, pem=0.0, similarity=21.07459]

6: 3536552 FC_Vorskla_Poltava
[pem_pos=0, final_scores=0.49601465,

```
final_local_scores=0.46398145, pem=0.0, similarity=17.935844]
7: 7010381 FC_Nyva_Ternopil
    [pem_pos=0, final_scores=0.36167216,
final_local_scores=0.35385877, pem=0.0, similarity=14.649565]
8: 7011579 FC_Torpedo_Zaporizhya
    [pem_pos=0, final_scores=0.4028157,
final_local_scores=0.4290874, pem=0.0, similarity=16.894537]
9: 861305 FC_Shakhtar_Donetsk
    [pem_pos=0, final_scores=0.40089065,
final_local_scores=0.34814233, pem=0.0, similarity=14.478974]
10: 4388256 FC_Karpaty_Lviv
    [pem_pos=0, final_scores=0.47307664,
final_local_scores=0.43309915, pem=0.0, similarity=17.014256]
11: 3435061 FC_Chornomorets_Odesa
    [pem_pos=0, final_scores=0.6224071,
final_local_scores=0.59530735, pem=0.0, similarity=21.854872]
12: 3517746 FC_Metalurh_Zaporizhya
    [pem_pos=0, final_scores=0.46913737,
final_local_scores=0.43266642, pem=0.0, similarity=17.001343]
13: 895551 FC_Dnipro_Dnipropetrovsk
    [pem_pos=0, final_scores=0.3764506,
final_local_scores=0.32326177, pem=0.0, similarity=13.736489]
14: 5056113 FC_CSKA_Kyiv
    [pem_pos=0, final_scores=0.31756014,
final_local_scores=0.32747674, pem=0.0, similarity=13.862272]
15: 889025 FC_Dynamo_Kyiv
    [pem_pos=5, final_scores=0.21191539,
final_local_scores=0.13458282, pem=-3.3873816,
similarity=26.527039]
16: 3536552 FC_Vorskla_Poltava
    [pem_pos=0, final_scores=0.45355076,
final_local_scores=0.4175961, pem=0.0, similarity=16.551613]
17: 895551 FC_Dnipro_Dnipropetrovsk
    [pem_pos=0, final_scores=0.35884625,
final_local_scores=0.30403167, pem=-0.5291793,
similarity=16.040384]
18: 3435061 FC_Chornomorets_Odesa
    [pem_pos=0, final_scores=0.5279466,
final_local_scores=0.4921236, pem=0.0, similarity=18.775663]
19: 861305 FC_Shakhtar_Donetsk
    [pem_pos=0, final_scores=0.39856577,
final_local_scores=0.34560272, pem=0.0, similarity=14.403188]
20: 17864564 FK_Metallurg_Skopje
    [pem_pos=0, final_scores=0.34980708,
final_local_scores=0.38687098, pem=0.0, similarity=15.634715]
21: 4388256 FC_Karpaty_Lviv
    [pem_pos=0, final_scores=0.46629578,
final_local_scores=0.42569214, pem=0.0, similarity=16.793215]
22: 1443555 FC_Torpedo_Moscow
    [pem_pos=3, final_scores=0.37782425,
final_local_scores=0.33673075, pem=-2.5035305,
similarity=27.753017]
```

```
23: 3325822 SC_Tavriya_Simferopol
    [pem_pos=1, final_scores=0.4012453,
final_local_scores=0.36146706, pem=-0.48393673,
similarity=17.508335]
24: 7010381 FC_Nyva_Ternopil
    [pem_pos=0, final_scores=0.37044173,
final_local_scores=0.3634382, pem=0.0, similarity=14.935433]
25: 3800152 FC_Kryvbas_Kryvyi_Rih
    [pem_pos=1, final_scores=0.46244735,
final_local_scores=0.45992577, pem=-0.13353139,
similarity=18.54098]
26: 7009763 MFC_Kremin_Kremenchuk
    [pem_pos=0, final_scores=0.38496315,
final_local_scores=0.37921053, pem=0.0, similarity=15.406112]
```

Sample: b'421_SWIMMING'

SWIMMING - POPOV IN `SERIOUS CONDITION ' AFTER STABBING . [1
1 MOSCOW 1] 1996-08-26 [2 2 Russian 2] double [3 3 Olympic 3]
swimming champion [4 4 Alexander Popov 4] was in a serious
condition on Monday after being stabbed on a [5 5 Moscow 5]
street . A doctor said it was too early to say whether
Popov , the only man to retain the [6 6 Olympic 6] 50 and 100
metres freestyle titles , would return to top-level sport . "
His condition is serious , " said Rimma Maslova , deputy
chief doctor of Hospital No 31 in the [7 7 Russian 7]
capital . " But he is conscious and is talking and smiling .
" Maslova told [8 8 Reuters 8] she was not an expert in
sports medicine , but said it was too early to judge Popov 's
chances of returning to competitive swimming . Popov , who
won gold in the 50 and 100 metres freestyle at the recent [9 9]
Atlanta Olympics 9] , was stabbed in the abdomen late on
Saturday after an argument with a group of roadside
watermelon sellers in south-west [10 10 Moscow 10] . Maslova
said the wound had affected a lung and a kidney . Doctors
operated on Popov , 24 , for three hours . Popov told [11 11 N
TV 11] television on Sunday he was in no danger and promised
he would be back in the pool shortly . " There 's no need to
worry . We 're going to be walking soon -- and swimming , "
he insisted cheerfully from his bed in the intensive care
unit . [12 12 Interfax 12] news agency said police had
detained one of the attackers . It said the row started when
Popov and a group of his friends were returning from a
party . Vitaly Smirnov , president of the Russian National
Olympic Committee , said President [13 13 Boris Yeltsin 13]
had given the swimmer [14 14 Russia 14] 's top award for his [
15 15 Olympic 15] performance . " I am not a doctor but I
think he is doing all right , " said 1 Smirnov . 2 Smirnov
said the [16 16 Olympic Committee 16] might ask the
government to take measures to protect the country 's best
athletes , some of whom have already chosen to live abroad
for fear of a surge in crime in [17 17 post-Soviet 17] [18 18
Russia 18] .

1: 19004 Moscow
2: 25391 Russia
3: 27566 Summer_Olympic_Games
4: 915621 Alexander_Popov_(swimmer)
5: 19004 Moscow
6: 27566 Summer_Olympic_Games
7: 25391 Russia
8: 18998750 Reuters
9: 72311 1996_Summer_Olympics
10: 19004 Moscow
11: 998231 NTV_(Russia)
12: 4921404 Interfax
13: 23278520 Boris_Yeltsin
14: 25391 Russia
15: 27566 Summer_Olympic_Games
16: 1217591 National_Olympic_Committee
17: 1609192 Post-Soviet_states
18: 25391 Russia
1: 18898893 Vitali_Smirnov
[winner: 19556446
Dmitry_Nikolayevich_Smirnov_(footballer), pem_pos=0,
final_scores=-0.48904234, final_local_scores=-0.38518336,
pem=-1.2714295, similarity=-0.4906953]
RECALL MISS: 18898893 Vitali_Smirnov]
2: 18898893 Vitali_Smirnov
[winner: 327841 Vladimir_Smirnov_(fencer), pem_pos=3,
final_scores=-0.44940126, final_local_scores=-0.39424106,
pem=-1.9943613, similarity=3.17042]
RECALL MISS: 18898893 Vitali_Smirnov]
1: 19004 Moscow
[pem_pos=0, final_scores=0.74007016,
final_local_scores=0.72046006, pem=-0.106582455,
similarity=26.16929]
2: 25391 Russia
[pem_pos=1, final_scores=0.7667355,
final_local_scores=0.7640327, pem=-1.0515022,
similarity=32.608204]
3: 27566 Summer_Olympic_Games
[pem_pos=2, final_scores=0.34034312,
final_local_scores=0.33275992, pem=-3.2532303,
similarity=31.711504]
4: 915621 Alexander_Popov_(swimmer)
[pem_pos=1, final_scores=0.4514464,
final_local_scores=0.52813375, pem=-0.49902612,
similarity=22.564058]
5: 19004 Moscow
[pem_pos=0, final_scores=0.68453616,
final_local_scores=0.6597976, pem=-0.106582455,
similarity=24.359001]
6: 27566 Summer_Olympic_Games
[pem_pos=2, final_scores=0.29350716,
final_local_scores=0.28159875, pem=-3.2532303,

```
similarity=30.184755]
7: 25391 Russia
    [pem_pos=1, final_scores=0.7930495,
final_local_scores=0.7927768, pem=-1.0515022,
similarity=33.46599]
8: 18998750 Reuters
    [pem_pos=0, final_scores=0.3538171,
final_local_scores=0.40765768, pem=0.0, similarity=16.255032]
9: 72311 1996_Summer_Olympics
    [pem_pos=0, final_scores=0.32893556,
final_local_scores=0.3172921, pem=0.0, similarity=13.558342]
10: 19004 Moscow
    [pem_pos=0, final_scores=0.6702407,
final_local_scores=0.644182, pem=-0.106582455,
similarity=23.893002]
11: 998231 NTV_(Russia)
    [pem_pos=1, final_scores=0.3251503,
final_local_scores=0.37823862, pem=-1.0768106,
similarity=21.232971]
12: 4921404 Interfax
    [pem_pos=0, final_scores=0.5328288,
final_local_scores=0.61571884, pem=0.0, similarity=22.463991]
13: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.634373,
final_local_scores=0.6887572, pem=0.0, similarity=24.6436]
14: 25391 Russia
    [pem_pos=0, final_scores=0.86895996,
final_local_scores=0.87569743, pem=-0.20299491,
similarity=31.326189]
15: 27566 Summer_Olympic_Games
    [pem_pos=2, final_scores=0.26095974,
final_local_scores=0.24604565, pem=-3.2532303,
similarity=29.12378]
16: 1217591 National_Olympic_Committee
    [pem_pos=2, final_scores=0.19423793,
final_local_scores=0.22740892, pem=-1.2872697,
similarity=17.876421]
17: 1609192 Post-Soviet_states
    [pem_pos=0, final_scores=0.4219584,
final_local_scores=0.43209618, pem=0.0, similarity=16.984325]
18: 25391 Russia
    [pem_pos=0, final_scores=0.7953039,
final_local_scores=0.7952393, pem=-0.20299491,
similarity=28.925158]
```

Sample: b'422_SOCCER'
SOCCER - [1 1 SLOVAK 1] FIRST DIVISION RESULTS / STANDINGS .
[2 2 BRATISLAVA 2] 1996-08-26 Results of [3 3 Slovak 3] first
division soccer matches at the weekend : [4 4 Inter Bratislava
4] 0 [5 5 Slovan Bratislava 5] 2 Chemlon Humenne 0 [6 6 Tatran
Presov 6] 1 Artmedia Petralka 0 JAS Bardejov 0 [7 7 DAC Duna
jska Streda 7] 1 [8 8 Spartak Trnava 8] 3 [9 9 Dukla Banska By

strica 9] 3 [10 10 FC Nitra 10] 0 [11 11 MSK Zilina 11] 0 [12 12 FC Košice 12] 2 Petrimex Prievidza 2 [13 13 FC Rimavská Sobota 13] 0 [14 14 Lokomotiva Košice 14] 2 [15 15 Kerametal Dubnica 15] 0 Standings (tabulate under games played , won , drawn , lost , goals for , goals against , points) : [16 16 Tatran Presov 16] 4 3 1 0 5 0 10 [17 17 Dukla Banská Bystrica 17] 4 3 0 1 7 2 9 [18 18 ŠK Slovan Bratislava 18] 4 3 0 1 7 2 9 Petrimex Prievidza 4 3 0 1 4 2 9 [19 19 Spartak Trnava 19] 4 2 2 0 10 5 8 [20 20 FC Košice 20] 4 2 2 0 6 3 8 Artmedia Petrzalka 4 1 3 0 1 0 6 [21 21 DAC Dunajská Streda 21] 4 2 0 2 5 6 6 [22 22 FC Rimavská Sobota 22] 4 2 0 2 4 5 6 JAS Bardejov 4 1 2 1 2 2 5 Chemlon Humenne 4 1 1 2 1 2 4 [23 23 Inter Bratislava 23] 4 1 1 2 4 6 4 [24 24 Lokomotiva Košice 24] 4 1 1 2 3 5 4 [25 25 Kerametal Dubnica 25] 4 0 1 3 3 9 1 [26 26 FC Nitra 26] 4 0 0 4 1 8 0 [27 27 MSK Zilina 27] 4 0 0 4 0 6 0

1: 26830 Slovakia
2: 18933194 Bratislava
3: 26830 Slovakia
4: 4414525 FK_Inter_Bratislava
5: 1664787 ŠK_Slovan_Bratislava
6: 6155772 1._FC_Tatran_Prešov
7: 6157701 FK_DAC_1904_Dunajská_Streda
8: 4403792 FC_Spartak_Trnava
9: 827800 FK_Dukla_Banská_Bystrica
10: 4783760 FC_Nitra
11: 861349 MŠK_Zilina
12: 5508993 MFK_Košice
13: 6155152 MŠK_Rimavská_Sobota
14: 16963970 FC_Lokomotíva_Košice
15: 4414792 MFK_Dubnica
16: 6155772 1._FC_Tatran_Prešov
17: 827800 FK_Dukla_Banská_Bystrica
18: 1664787 ŠK_Slovan_Bratislava
19: 4403792 FC_Spartak_Trnava
20: 5508993 MFK_Košice
21: 6157701 FK_DAC_1904_Dunajská_Streda
22: 6155152 MŠK_Rimavská_Sobota
23: 4414525 FK_Inter_Bratislava
24: 16963970 FC_Lokomotíva_Košice
25: 4414792 MFK_Dubnica
26: 4783760 FC_Nitra
27: 861349 MŠK_Zilina

1: 26830 Slovakia
[pem_pos=1, final_scores=0.53827304, final_local_scores=0.55950326, pem=-1.0527612, similarity=26.511484]
2: 18933194 Bratislava
[pem_pos=0, final_scores=0.43645805, final_local_scores=0.48603892, pem=0.0, similarity=18.594084]
3: 26830 Slovakia
[pem_pos=1, final_scores=0.4852187,

```
final_local_scores=0.5015494, pem=-1.0527612,
similarity=24.782026]
4: 4414525 FK_Inter_Bratislava
    [pem_pos=0, final_scores=0.4070478,
final_local_scores=0.3817004, pem=0.0, similarity=15.480413]
5: 1664787 ŠK_Slovan_Bratislava
    [pem_pos=0, final_scores=0.41080105,
final_local_scores=0.36084998, pem=0.0, similarity=14.858196]
6: 6155772 Ľ. FC_Tatran_Prešov
    [pem_pos=0, final_scores=0.5876673,
final_local_scores=0.57358867, pem=0.0, similarity=21.206743]
7: 6157701 FK_DAC_1904_Dunajská_Streda
    [pem_pos=0, final_scores=0.57169205,
final_local_scores=0.57334924, pem=0.0, similarity=21.199598]
8: 4403792 FC_Spartak_Trnava
    [pem_pos=0, final_scores=0.48398572,
final_local_scores=0.4405644, pem=0.0, similarity=17.237034]
9: 827800 FK_Dukla_Banská_Bystrica
    [pem_pos=0, final_scores=0.5245803,
final_local_scores=0.50302553, pem=0.0, similarity=19.100998]
10: 4783760 FC_Nitra
    [pem_pos=0, final_scores=0.45847815,
final_local_scores=0.43061298, pem=0.0, similarity=16.940063]
11: 861349 MŠK_Žilina
    [pem_pos=0, final_scores=0.4789394,
final_local_scores=0.45912665, pem=0.0, similarity=17.790968]
12: 5508993 MFK_Košice
    [pem_pos=0, final_scores=0.45919174,
final_local_scores=0.41194737, pem=0.0, similarity=16.383043]
13: 6155152 MŠK_Rimavská_Sobota
    [pem_pos=0, final_scores=0.4837973,
final_local_scores=0.504708, pem=0.0, similarity=19.151207]
14: 16963970 FC_Lokomotíva_Košice
    [pem_pos=0, final_scores=0.38794976,
final_local_scores=0.38653874, pem=0.0, similarity=15.624801]
15: 4414792 MFK_Dubnica
    [pem_pos=0, final_scores=0.50747204,
final_local_scores=0.502195, pem=0.0, similarity=19.076214]
16: 6155772 Ľ. FC_Tatran_Prešov
    [pem_pos=0, final_scores=0.44196063,
final_local_scores=0.41442627, pem=0.0, similarity=16.45702]
17: 827800 FK_Dukla_Banská_Bystrica
    [pem_pos=0, final_scores=0.51166075,
final_local_scores=0.48891288, pem=0.0, similarity=18.679848]
18: 1664787 ŠK_Slovan_Bratislava
    [pem_pos=0, final_scores=0.4536934,
final_local_scores=0.40770334, pem=0.0, similarity=16.256393]
19: 4403792 FC_Spartak_Trnava
    [pem_pos=0, final_scores=0.4894988,
final_local_scores=0.4465866, pem=0.0, similarity=17.416748]
20: 5508993 MFK_Košice
    [pem_pos=0, final_scores=0.43835497,
```

```

final_local_scores=0.38918632, pem=0.0, similarity=15.703809]
21: 6157701 FK_DAC_1904_Dunajská_Streda
    [pem_pos=0, final_scores=0.5528333,
final_local_scores=0.552749, pem=0.0, similarity=20.584845]
22: 6155152 MŠK_Rimavská_Sobota
    [pem_pos=0, final_scores=0.45280713,
final_local_scores=0.47085595, pem=0.0, similarity=18.140993]
23: 4414525 FK_Inter_Bratislava
    [pem_pos=0, final_scores=0.41466528,
final_local_scores=0.39002138, pem=0.0, similarity=15.728728]
24: 16963970 FC_Lokomotíva_Košice
    [pem_pos=0, final_scores=0.38793397,
final_local_scores=0.38652146, pem=0.0, similarity=15.624284]
25: 4414792 MFK_Dubnica
    [pem_pos=0, final_scores=0.47999227,
final_local_scores=0.4721775, pem=0.0, similarity=18.180431]
26: 4783760 FC_Nitra
    [pem_pos=0, final_scores=0.48898846,
final_local_scores=0.46394086, pem=0.0, similarity=17.934633]
27: 861349 MŠK_Žilina
    [pem_pos=0, final_scores=0.45402223,
final_local_scores=0.43190837, pem=0.0, similarity=16.97872]

```

Sample: b'423_SOCCER'

```

SOCCER - [1 1 HUNGARY 1] FIRST DIVISION RESULTS /
STANDINGS . [2 2 BUDAPEST 2] 1996-08-26 [3 3 Hungarian 3]
first division soccer results and standings from weekend and
bank holiday : Stadler 0 Haladas 0 [4 4 MTK 4] 3 [5 5 Ferencv
aros 5] 0 Bekescsaba 0 [6 6 BVSC 6] 1 [7 7 Csepel 7] 1
Videoton(*) 1 [8 8 ZTE 8] 1 [9 9 Debrecen 9] 5 Siofok 0 [10 1
0 Ujpest 10] 2 Vac 0 [11 11 Vasas 11] 1 Kispest 3 Pecs 1
Standings ( tabulated under played , won , drawn , lost ,
goals for , goals against , points ) : 1. Ujpest TE 3 3 - -
10 2 9 2. [12 12 MTK 12] 3 3 - - 7 1 9 3. [13 13 BVSC 13] 3 2
1 - 6 2 7 4. [14 14 Debrecen 14] 3 2 - 1 10 4 6 5. Bekescsaba
3 2 - 1 6 2 6 6. [15 15 FTC 15] 3 2 - 1 8 7 6 7. Haladas 3 1
2 - 2 1 5 8. [16 16 Videoton 16] 3 1 1 1 7 5 4 9. [17 17 Vasas
17] 3 1 1 1 3 3 4 10. Kispest 3 1 1 1 6 7 4 11. [18 18 Gyor 18
1] 3 1 1 1 3 5 4 12. [19 19 Csepel 19] 3 - 3 - 3 3 3 13. Pecs
3 1 - 2 3 5 3 14. [20 20 ZTE 20] 3 1 - 2 3 10 3 15. Stadler
FC 3 - 1 2 2 6 1 16. III.ker.TVE 3 - 1 2 2 7 1 17. Siofok 3 -
- 3 2 7 0 18. Vac 3 - - 3 2 8 0 *Name of Parmalat / Fehervar
FC has been changed to [21 21 Videoton 21] . -- [22 22 Budape
st 22] newsroom , +361 266 2410

```

- 1: 13275 Hungary
- 2: 36787 Budapest
- 3: 13275 Hungary
- 4: 3205346 MTK_Budapest_FC
- 5: 861261 Ferencvárosi_TC
- 6: 5798281 BVSC_Budapest
- 7: 9837826 Csepel_SC
- 8: 1688457 Zalaegerszegi_TE

9: 32821489 Debreceni_VSC
10: 896074 Újpest_FC
11: 4719597 Vasas_SC
12: 3205346 MTK_Budapest_FC
13: 5798281 BVSC_Budapest
14: 32821489 Debreceni_VSC
15: 861261 Ferencvárosi_TC
16: 30865546 Videoton_FC
17: 4719597 Vasas_SC
18: 5798733 Győri_ETO_FC
19: 9837826 Csepel_SC
20: 1688457 Zalaegerszegi_TE
21: 30865546 Videoton_FC
22: 36787 Budapest
1: 13275 Hungary
[pem_pos=0, final_scores=0.53915185,
final_local_scores=0.5611612, pem=-0.32577527,
similarity=22.6075]
2: 36787 Budapest
[pem_pos=0, final_scores=0.5405031,
final_local_scores=0.56524014, pem=-0.030229695,
similarity=21.122]
3: 13275 Hungary
[pem_pos=1, final_scores=0.60451907,
final_local_scores=0.6325649, pem=-1.0319586,
similarity=28.578663]
4: 3205346 MTK_Budapest_FC
[pem_pos=1, final_scores=0.45201653,
final_local_scores=0.402928, pem=-0.51316404,
similarity=18.904554]
5: 861261 Ferencvárosi_TC
[pem_pos=0, final_scores=0.46490777,
final_local_scores=0.41509992, pem=0.0, similarity=16.477123]
6: 5798281 BVSC_Budapest
[pem_pos=0, final_scores=0.3335771,
final_local_scores=0.333718, pem=0.0, similarity=14.048524]
7: 9837826 Csepel_SC
[pem_pos=1, final_scores=0.3802113,
final_local_scores=0.38571388, pem=-0.58512866,
similarity=18.782206]
8: 1688457 Zalaegerszegi_TE
[pem_pos=1, final_scores=0.53956956,
final_local_scores=0.5307153, pem=-0.6000891,
similarity=23.190693]
9: 32821489 Debreceni_VSC
[pem_pos=1, final_scores=0.39428812,
final_local_scores=0.35473067, pem=-1.1268859,
similarity=20.803764]
10: 896074 Újpest_FC
[pem_pos=0, final_scores=0.46015495,
final_local_scores=0.40909147, pem=0.0, similarity=16.29782]
11: 4719597 Vasas_SC

```

[pem_pos=0, final_scores=0.41174048,
final_local_scores=0.3646767, pem=0.0, similarity=14.972394]
12: 3205346 MTK_Budapest_FC
[pem_pos=1, final_scores=0.6054441,
final_local_scores=0.5705242, pem=-0.51316404,
similarity=23.90596]
13: 5798281 BVSC_Budapest
[pem_pos=0, final_scores=0.40863353,
final_local_scores=0.41570568, pem=0.0, similarity=16.4952]
14: 32821489 Debreceni_VSC
[pem_pos=1, final_scores=0.37692183,
final_local_scores=0.3357606, pem=-1.1268859,
similarity=20.23766]
15: 861261 Ferencvárosi_TC
[pem_pos=1, final_scores=0.65173143,
final_local_scores=0.61917645, pem=-2.1957188,
similarity=34.507835]
16: 30865546 Videoton_FC
[pem_pos=0, final_scores=0.62074035,
final_local_scores=0.61062986, pem=0.0, similarity=22.312126]
17: 4719597 Vasas_SC
[pem_pos=0, final_scores=0.44508046,
final_local_scores=0.4010955, pem=0.0, similarity=16.059202]
18: 5798733 Györi_ET0_FC
[pem_pos=1, final_scores=0.43007952,
final_local_scores=0.40366596, pem=-1.377265,
similarity=23.625692]
19: 9837826 Csepel_SC
[pem_pos=1, final_scores=0.3724057,
final_local_scores=0.3771875, pem=-0.58512866,
similarity=18.527761]
20: 1688457 Zalaegerszegi_TE
[pem_pos=1, final_scores=0.43374562,
final_local_scores=0.41511875, pem=-0.6000891,
similarity=19.741062]
21: 30865546 Videoton_FC
[pem_pos=0, final_scores=0.40527397,
final_local_scores=0.37526554, pem=0.0, similarity=15.288385]
22: 36787 Budapest
[pem_pos=0, final_scores=0.42993683,
final_local_scores=0.44446325, pem=-0.030229695,
similarity=17.517776]

```

Sample: b'424_SOCCER'

SOCCER - [1 1 CZECH 1] FIRST DIVISION RESULTS / STANDINGS . [
2 2 PRAGUE 2] 1996-08-26 Results of the [3 3 Czech Republic 3]
's first division soccer matches at the weekend : [4 4 Petra D
rnovice 4] 1 [5 5 Slovan Liberec 5] 3 [6 6 SK Slavia Praha 6]
3 [7 7 SK Ceske Budejovice 7] 0 [8 8 FK Jablonec 8] 3 [9 9 Vi
ktorija Zizkov 9] 1 [10 10 Banik Ostrava 10] 3 [11 11 FK Teplice 11]
1 [12 12 Boby Brno 12] 1 [13 13 Sigma Olomouc 13] 0 [14
14 FC Bohemians 14] 0 FC Karvina 2 [15 15 SK Hradec Kralove 15]

] 0 [16 16 Kaucuk Opava 16] 0 Playing Monday : [17 17 Viktoria Plzen 17] v [18 18 AC Sparta Praha 18] Standings (tabulate under games played , won , drawn , lost , goals for , against , points) : [19 19 Boby Brno 19] 3 3 0 0 5 2 9 [20 20 Banik Ostrava 20] 3 2 0 1 7 3 6 [21 21 FK Jablonec 21] 3 2 0 1 5 2 6 [22 22 SK Slavia Praha 22] 3 1 2 0 6 3 5 [23 23 Kaucuk Opava 23] 3 1 2 0 2 1 5 [24 24 Sigma Olomouc 24] 3 1 1 1 1 6 3 4 [25 25 Petra Drnovice 25] 3 1 1 1 7 5 4 [26 26 Slovan Liberec 26] 3 1 1 1 5 4 4 [27 27 FK Teplice 27] 3 1 1 1 3 4 4 FC Karvina 3 1 1 1 3 5 4 [28 28 SK Ceske Budejovice 28] 3 1 1 1 3 5 4 [29 29 Viktoria Plzen 29] 2 0 2 0 2 2 2 [30 30 AC Sparta Praha 30] 2 0 1 1 3 4 1 [31 31 FC Bohemians 31] 3 0 1 2 1 4 1 [32 32 Viktoria Zizkov 32] 3 0 1 2 3 8 1 [33 33 SK Hradec Králové 33] 3 0 1 2 1 7 1
1: 5321 Czech_Republic
2: 23844 Prague
3: 5321 Czech_Republic
4: 1650918 FK_Drnovice
5: 1650901 FC_Slovan_Liberec
6: 896034 SK_Slavia_Prague
7: 4746509 SK_Dynamo_České_Budějovice
8: 4991040 FK_Baumit_Jablonec
9: 3752998 FK_Viktoria_Žižkov
10: 1138124 FC_Baník_Ostrava
11: 1650910 FK_Teplice
12: 3310797 FC_Zbrojovka_Brno
13: 895511 SK_Sigma_Olomouc
14: 1004409 Bohemians_1905
15: 5347958 FC_Hradec_Králové
16: 8146110 SFC_Opava
17: 4036422 FC_Viktoria_Plzeň
18: 861331 AC_Sparta_Prague
19: 3310797 FC_Zbrojovka_Brno
20: 1138124 FC_Baník_Ostrava
21: 4991040 FK_Baumit_Jablonec
22: 896034 SK_Slavia_Prague
23: 8146110 SFC_Opava
24: 895511 SK_Sigma_Olomouc
25: 1650918 FK_Drnovice
26: 1650901 FC_Slovan_Liberec
27: 1650910 FK_Teplice
28: 4746509 SK_Dynamo_České_Budějovice
29: 4036422 FC_Viktoria_Plzeň
30: 861331 AC_Sparta_Prague
31: 1004409 Bohemians_1905
32: 3752998 FK_Viktoria_Žižkov
33: 5347958 FC_Hradec_Králové
1: 5321 Czech_Republic
[pem_pos=1, final_scores=0.5325442, final_local_scores=0.575052, pem=-1.208016, similarity=27.819792]
2: 23844 Prague

[pem_pos=0, final_scores=0.4123147,
final_local_scores=0.46595204, pem=-0.05392049,
similarity=18.287878]
3: 5321 Czech_Republic
[pem_pos=0, final_scores=0.43159872,
final_local_scores=0.46478444, pem=-0.27794555,
similarity=19.47132]
4: 1650918 FK_Drnovice
[pem_pos=0, final_scores=0.5358195,
final_local_scores=0.53149074, pem=0.0, similarity=19.950457]
5: 1650901 FC_Slovan_Liberec
[pem_pos=0, final_scores=0.43907076,
final_local_scores=0.39892566, pem=0.0, similarity=15.994451]
6: 896034 SK_Slavia_Prague
[pem_pos=0, final_scores=0.5653311,
final_local_scores=0.52261984, pem=0.0, similarity=19.685732]
7: 4746509 SK_Dynamo_České_Budějovice
[pem_pos=0, final_scores=0.5454239,
final_local_scores=0.5349406, pem=0.0, similarity=20.053408]
8: 4991040 FK_Baumit_Jablonec
[pem_pos=0, final_scores=0.5670052,
final_local_scores=0.5349361, pem=0.0, similarity=20.053272]
9: 3752998 FK_Viktoria_Žižkov
[pem_pos=0, final_scores=0.884396,
final_local_scores=0.89000815, pem=0.0, similarity=30.649332]
10: 1138124 FC_Baník_Ostrava
[pem_pos=0, final_scores=0.61860013,
final_local_scores=0.5891212, pem=0.0, similarity=21.670265]
11: 1650910 FK_Teplice
[pem_pos=0, final_scores=0.5588035,
final_local_scores=0.530438, pem=0.0, similarity=19.91904]
12: 3310797 FC_Zbrojovka_Brno
[pem_pos=0, final_scores=0.74833256,
final_local_scores=0.7328468, pem=0.0, similarity=25.959324]
13: 895511 SK_Sigma_Olomouc
[pem_pos=0, final_scores=0.61141056,
final_local_scores=0.5922304, pem=0.0, similarity=21.763048]
14: 1004409 Bohemians_1905
[pem_pos=1, final_scores=0.41086966,
final_local_scores=0.38072085, pem=-0.44183278,
similarity=17.85394]
15: 5347958 FC_Hradec_Králové
[pem_pos=0, final_scores=0.5654812,
final_local_scores=0.5582121, pem=0.0, similarity=20.747875]
16: 8146110 SFC_Opava
[pem_pos=0, final_scores=0.6013339,
final_local_scores=0.6198764, pem=0.0, similarity=22.58806]
17: 4036422 FC_Viktoria_Plzeň
[pem_pos=0, final_scores=0.59458005,
final_local_scores=0.5728767, pem=0.0, similarity=21.185495]
18: 861331 AC_Sparta_Prague
[pem_pos=0, final_scores=0.5303518,

```
final_local_scores=0.48893338, pem=0.0, similarity=18.68046]
19: 3310797 FC_Zbrojovka_Brno
    [pem_pos=0, final_scores=0.71534914,
final_local_scores=0.6968175, pem=0.0, similarity=24.884138]
20: 1138124 FC_Baník_Ostrava
    [pem_pos=0, final_scores=0.6188508,
final_local_scores=0.58939505, pem=0.0, similarity=21.678436]
21: 4991040 FK_Baumit_Jablonec
    [pem_pos=0, final_scores=0.5717645,
final_local_scores=0.54013485, pem=0.0, similarity=20.208414]
22: 896034 SK_Slavia_Prague
    [pem_pos=0, final_scores=0.5626737,
final_local_scores=0.51971704, pem=0.0, similarity=19.599106]
23: 8146110 SFC_Opava
    [pem_pos=0, final_scores=0.56772584,
final_local_scores=0.58316475, pem=0.0, similarity=21.492512]
24: 895511 SK_Sigma_Olomouc
    [pem_pos=0, final_scores=0.585849,
final_local_scores=0.5643083, pem=0.0, similarity=20.929798]
25: 1650918 FK_Drnovice
    [pem_pos=0, final_scores=0.59060067,
final_local_scores=0.5913309, pem=0.0, similarity=21.736206]
26: 1650901 FC_Slovan_Liberec
    [pem_pos=0, final_scores=0.41677588,
final_local_scores=0.37457192, pem=0.0, similarity=15.267686]
27: 1650910 FK_Teplice
    [pem_pos=0, final_scores=0.5517261,
final_local_scores=0.52270705, pem=0.0, similarity=19.688334]
28: 4746509 SK_Dynamo_České_Budějovice
    [pem_pos=0, final_scores=0.5270752,
final_local_scores=0.5148974, pem=0.0, similarity=19.455278]
29: 4036422 FC_Viktoria_Plzeň
    [pem_pos=0, final_scores=0.6265912,
final_local_scores=0.607844, pem=0.0, similarity=22.22899]
30: 861331 AC_Sparta_Prague
    [pem_pos=0, final_scores=0.46753073,
final_local_scores=0.42031085, pem=0.0, similarity=16.632627]
31: 1004409 Bohemians_1905
    [pem_pos=1, final_scores=0.36678547,
final_local_scores=0.3325656, pem=-0.44183278,
similarity=16.41689]
32: 3752998 FK_Viktoria_Žižkov
    [pem_pos=0, final_scores=0.8846869,
final_local_scores=0.89032584, pem=0.0, similarity=30.658812]
33: 5347958 FC_Hradec_Králové
    [pem_pos=0, final_scores=0.58830196,
final_local_scores=0.5831403, pem=0.0, similarity=21.491783]
```

Sample: b'425_SOCCER'

SOCCER - [1 1 MEXICAN 1] CHAMPIONSHIP RESULTS / STANDINGS . [
2 2 MEXICO CITY 2] 1996-08-26 Results of weekend matches in
the [3 3 Mexican 3] soccer championship : [4 4 Atlante 4] 1 [5

5 Atlas 5] 1 [6 6 Cruz Azul 6] 2 1 Leon 2 [7 7 Guadalajara 7]
 5 [8 8 America 8] 0 [9 9 Monterrey 9] 2 [10 10 Veracruz 10] 1
 [11 11 Pachuca 11] 3 [12 12 Toluca 12] 0 [13 13 Puebla 13] 2 [
 14 14 UNAM 14] 1 [1 2 Santos 1] 2 [15 15 Morelia 15] 1 [16 16
 UAG 16] 1 [17 17 Neza 17] 2 Standings (tabulated under
 played , won , drawn , lost , goals for , against , points)
 Group 1 [18 18 Puebla 18] 3 3 0 0 7 2 9 [19 19 Cruz Azul 19]
 3 2 1 0 7 3 7 [20 20 Atlante 20] 3 2 1 0 6 2 7 [21 21 Neza 21]
 3 1 0 2 2 7 3 [22 22 Veracruz 22] 3 0 1 2 2 6 1 Group 2 [23 23
 Necaxa 23] 3 1 1 1 6 4 4 [24 24 Pachuca 24] 3 1 1 1 6 7 4 3 Le
 on 3 0 3 0 3 3 3 [25 25 America 25] 3 1 0 2 5 7 3 [26 26 More
 lia 26] 3 0 1 2 3 8 1 Group 3 [27 27 Atlas 27] 3 2 1 0 7 2 7 [
 28 28 Guadalajara 28] 3 2 1 0 7 0 7 [29 29 Toluca 29] 3 1 0 2
 6 5 3 [30 30 UNAM 30] 3 0 0 3 2 6 0 Group 4 [31 31 Santos 31]
 3 3 0 0 4 1 9 [32 32 Monterrey 32] 4 1 1 2 2 5 4 [33 33 Celaya
 33] 2 0 2 0 1 1 2 [34 34 UAG 34] 3 0 0 3 1 8 0
 1: 3966054 Mexico
 2: 18987 Mexico_City
 3: 3966054 Mexico
 4: 7445919 C.F._Atlante
 5: 859774 Club_Atlas
 6: 847584 Cruz_Azul
 7: 20878980 C.D._Guadalajara
 8: 1025920 Club_America
 9: 547412 C.F._Monterrey
 10: 1193316 Tiburones_Rojos_de_Veracruz
 11: 17567911 C.F._Pachuca
 12: 1361541 Deportivo_Toluca_F.C.
 13: 1511135 Puebla_F.C.
 14: 553186 Club_Universidad_Nacional
 15: 1035876 Monarcas_Morelia
 16: 1511138 Estudiantes_Tecos
 17: 8619326 Toros_Neza
 18: 1511135 Puebla_F.C.
 19: 847584 Cruz_Azul
 20: 7445919 C.F._Atlante
 21: 8619326 Toros_Neza
 22: 1193316 Tiburones_Rojos_de_Veracruz
 23: 1221360 Club_Necaxa
 24: 17567911 C.F._Pachuca
 25: 1025920 Club_America
 26: 1035876 Monarcas_Morelia
 27: 859774 Club_Atlas
 28: 20878980 C.D._Guadalajara
 29: 1361541 Deportivo_Toluca_F.C.
 30: 553186 Club_Universidad_Nacional
 31: 187246 Santos_FC
 32: 547412 C.F._Monterrey
 33: 425696 Club_Celaya
 34: 1511138 Estudiantes_Tecos
 1: 4147317 Club_León
 [winner: 163447 Leon_Spinks, pem_pos=13,

```
final_scores=-0.57598084, final_local_scores=-0.5301797,
pem=-3.8139274, similarity=9.008824]
    RECALL MISS: 4147317 Club_León]
2: 1217587 Club_Santos_Laguna
    [winner: 187246 Santos_FC, pem_pos=0,
final_scores=0.32146293, final_local_scores=0.29191816,
pem=-0.81493, similarity=17.23285]
    gt: 1217587 Club_Santos_Laguna, pem_pos=3,
final_scores=0.00520657, final_local_scores=-0.08459693,
pem=-3.342929, similarity=19.74454 ]
3: 4147317 Club_León
    [winner: 163447 Leon_Spinks, pem_pos=13,
final_scores=-0.750262, final_local_scores=-0.7205554,
pem=-3.8139274, similarity=3.3276322]
    RECALL MISS: 4147317 Club_León]
1: 3966054 Mexico
    [pem_pos=0, final_scores=0.7296598,
final_local_scores=0.8116284, pem=-0.3968012,
similarity=30.46819]
2: 18987 Mexico_City
    [pem_pos=0, final_scores=0.42665285,
final_local_scores=0.4721895, pem=-0.028392298,
similarity=18.335192]
3: 3966054 Mexico
    [pem_pos=0, final_scores=0.57914495,
final_local_scores=0.6472138, pem=-0.3968012,
similarity=25.561731]
4: 7445919 C.F._Atlante
    [pem_pos=0, final_scores=0.5222098,
final_local_scores=0.4667077, pem=0.0, similarity=18.017202]
5: 859774 Club_Atlas
    [pem_pos=2, final_scores=0.8235367,
final_local_scores=0.8025817, pem=-1.9620275,
similarity=38.710163]
6: 847584 Cruz_Azul
    [pem_pos=0, final_scores=0.37467867,
final_local_scores=0.313028, pem=0.0, similarity=13.431093]
7: 20878980 C.D._Guadalajara
    [pem_pos=1, final_scores=0.5080916,
final_local_scores=0.45890182, pem=-1.5660818,
similarity=26.300854]
8: 1025920 Club_América
    [pem_pos=13, final_scores=0.62560695,
final_local_scores=0.57830185, pem=-4.222916,
similarity=44.312267]
9: 547412 C.F._Monterrey
    [pem_pos=1, final_scores=0.495098,
final_local_scores=0.43582112, pem=-1.5559204,
similarity=25.556822]
10: 1193316 Tiburones_Rojos_de_Veracruz
    [pem_pos=2, final_scores=0.7483395,
final_local_scores=0.72135186, pem=-1.6309961,
```

```
similarity=34.4859]
11: 17567911 C.F._Pachuca
    [pem_pos=0, final_scores=0.6961635,
final_local_scores=0.6615857, pem=0.0, similarity=23.83275]
12: 1361541 Deportivo_Toluca_F.C.
    [pem_pos=1, final_scores=0.45496696,
final_local_scores=0.39709383, pem=-0.83415943,
similarity=20.476072]
13: 1511135 Puebla_F.C.
    [pem_pos=2, final_scores=0.48104745,
final_local_scores=0.425022, pem=-2.2396874,
similarity=28.952986]
14: 553186 Club_Universidad_Nacional
    [pem_pos=1, final_scores=0.5608265,
final_local_scores=0.517404, pem=-0.5901791,
similarity=22.739567]
15: 1035876 Monarcas_Morelia
    [pem_pos=1, final_scores=0.3814234,
final_local_scores=0.32643718, pem=-0.7843081,
similarity=18.09644]
16: 1511138 Estudiantes_Tecos
    [pem_pos=0, final_scores=0.48255038,
final_local_scores=0.43375832, pem=-0.022593744,
similarity=17.156796]
17: 8619326 Toros_Neza
    [pem_pos=0, final_scores=0.43356866,
final_local_scores=0.4185292, pem=0.0, similarity=16.57946]
18: 1511135 Puebla_F.C.
    [pem_pos=2, final_scores=0.43940192,
final_local_scores=0.37953067, pem=-2.2396874,
similarity=27.595434]
19: 847584 Cruz_Azul
    [pem_pos=0, final_scores=0.35691744,
final_local_scores=0.29362655, pem=0.0, similarity=12.852115]
20: 7445919 C.F._Atlante
    [pem_pos=0, final_scores=0.4558862,
final_local_scores=0.39425915, pem=0.0, similarity=15.855192]
21: 8619326 Toros_Neza
    [pem_pos=0, final_scores=0.46275455,
final_local_scores=0.45041043, pem=0.0, similarity=17.530859]
22: 1193316 Tiburones_Rojos_de_Veracruz
    [pem_pos=2, final_scores=0.48121542,
final_local_scores=0.4295594, pem=-1.6309961,
similarity=25.778233]
23: 1221360 Club_Necaxa
    [pem_pos=0, final_scores=0.39370555,
final_local_scores=0.3336334, pem=0.0, similarity=14.045999]
24: 17567911 C.F._Pachuca
    [pem_pos=0, final_scores=0.64897084,
final_local_scores=0.61003494, pem=0.0, similarity=22.294373]
25: 1025920 Club_América
    [pem_pos=13, final_scores=0.5642454,
```

```
final_local_scores=0.5112737, pem=-4.222916,
similarity=42.31201]
26: 1035876 Monarcas_Morelia
    [pem_pos=1, final_scores=0.48046613,
final_local_scores=0.43462628, pem=-0.7843081,
similarity=21.325018]
27: 859774 Club_Atlas
    [pem_pos=2, final_scores=0.7230791,
final_local_scores=0.69284695, pem=-1.9620275,
similarity=35.43546]
28: 20878980 C.D._Guadalajara
    [pem_pos=1, final_scores=0.4962418,
final_local_scores=0.44595772, pem=-1.5660818,
similarity=25.914577]
29: 1361541 Deportivo_Toluca_F.C.
    [pem_pos=1, final_scores=0.4092304,
final_local_scores=0.34713367, pem=-0.83415943,
similarity=18.985163]
30: 553186 Club_Universidad_Nacional
    [pem_pos=1, final_scores=0.57222205,
final_local_scores=0.529852, pem=-0.5901791,
similarity=23.111038]
31: 187246 Santos_FC
    [pem_pos=0, final_scores=0.27473062,
final_local_scores=0.24087018, pem=-0.81493,
similarity=15.709476]
32: 547412 C.F._Monterrey
    [pem_pos=1, final_scores=0.47519237,
final_local_scores=0.41407722, pem=-1.5559204,
similarity=24.90794]
33: 425696 Club_Celaya
    [pem_pos=1, final_scores=0.27387816,
final_local_scores=0.22879368, pem=-1.2101505,
similarity=17.49836]
34: 1511138 Estudiantes_Tecos
    [pem_pos=0, final_scores=0.51095456,
final_local_scores=0.46478558, pem=-0.022593744,
similarity=18.08271]
1: 187246 Santos_FC
    [pem_pos=0, final_scores=0.32146293,
final_local_scores=0.29191816, pem=-0.81493,
similarity=17.23285]
```

Sample: b'426_SOCCER'
SOCCER - PLAYERS LEAVE MATCH EARLY TO CATCH PLANE . Brian Homewood [1 1 RIO DE JANEIRO 1] 1996-08-26 Two key players left a [2 2 Brazilian 2] championship match early on Sunday because they had to catch a plane to [3 3 Russia 3] to play with the national team . [4 4 Sao Paulo 4] midfielder Andre and [5 5 Santos 5] defender [6 6 Narciso 6] were both substituted during their teams ' game , taken to [7 7 Sao Paulo 7] airport and flown to [8 8 Rio de Janeiro 8] in a

private jet chartered by the [9 9 Brazilian Football Confederation 9] ([10 10 CBF 10]) . At [11 11 Rio 11] , they joined up with the national team squad for the journey to [12 12 Moscow 12] , where [13 13 Brazil 13] will face [14 14 Russia 14] in a friendly international on Wednesday . The problem arose because the Sao Paulo-Santos clash was selected as the day 's televised live match , forcing it to be put back three hours from the usual kickoff time . [15 15 Santos 15] suffered more from their loss as [16 16 Narciso 16] 's replacement 1 Jean gave away a penalty from which [17 17 Sao Paulo 17] scored the decisive goal in a 2-1 win . [18 18 Sao Paulo 18] lead the first stage of the championship on goal difference from surprise package [19 19 Juventude 19] , who beat [20 20 Internacional 20] 2-1 . [21 21 Corinthians 21] , who played in a tournament in [22 22 Spain 22] last week , also faced a plane marathon as they attempted to keep up with a hectic fixture list . They were due to leave [23 23 Spain 23] Monday night , arrive in [24 24 Sao Paulo 24] on Tuesday morning , catch another plane to the southern city of [25 25 Curitiba 25] one hour later and then play away to [26 26 Atletico Paranaense 26] in the [27 27 Brazilian 27] championship the same evening . [28 28 Botafogo 28] striker [29 29 Tullio 29] , who was overlooked by Zagalo for the tour which also features a game away to the [30 30 Netherlands 30] on Sunday , scored his third goal in three games as the defending champions beat [31 31 Bahia 31] 2-1 away . [32 32 Tullio 32] , who has been top-scorer in the competition for the last two seasons , has been struggling against injury for most of this year .

- 1: 25936 Rio_de_Janeiro
- 2: 3383 Brazil
- 3: 25391 Russia
- 4: 187274 São_Paulo_FC
- 5: 187246 Santos_FC
- 6: 6189919 Narciso_dos_Santos
- 7: 390875 São_Paulo
- 8: 25936 Rio_de_Janeiro
- 9: 681926 Brazilian_Football_Confederation
- 10: 681926 Brazilian_Football_Confederation
- 11: 25936 Rio_de_Janeiro
- 12: 19004 Moscow
- 13: 149286 Brazil_national_football_team
- 14: 25127146 Russia_national_football_team
- 15: 187246 Santos_FC
- 16: 6189919 Narciso_dos_Santos
- 17: 187274 São_Paulo_FC
- 18: 187274 São_Paulo_FC
- 19: 1577507 Esporte_Clube_Juventude
- 20: 187291 Sport_Club_Internacional
- 21: 187241 Sport_Club_Corinthians_Paulista
- 22: 26667 Spain
- 23: 26667 Spain
- 24: 390875 São_Paulo

```
25: 439002 Curitiba
26: 187346 Clube_Atlético_Paranaense
27: 3383 Brazil
28: 187238 Botafogo_de_Futebol_e_Regatas
29: 2335808 Túlio_Maravilha
30: 9647657 Netherlands_national_football_team
31: 187328 Esporte_Clube_Bahia
32: 2335808 Túlio_Maravilha
1: 8817274 Jean_Carlo_Witte
    [winner: 19802445 Jean_Narde, pem_pos=14,
final_scores=-0.6312681, final_local_scores=-0.6015159,
pem=-3.838211, similarity=7.0120664]
    RECALL MISS: 8817274 Jean_Carlo_Witte]
1: 25936 Rio_de_Janeiro
    [pem_pos=0, final_scores=0.47740102,
final_local_scores=0.45170265, pem=0.0, similarity=17.569422]
2: 3383 Brazil
    [pem_pos=0, final_scores=0.7332188,
final_local_scores=0.7599113, pem=-0.2733083,
similarity=28.253275]
3: 25391 Russia
    [pem_pos=0, final_scores=0.57733494,
final_local_scores=0.66425246, pem=-0.20299491,
similarity=25.016249]
4: 187274 São_Paulo_FC
    [pem_pos=2, final_scores=0.43576425,
final_local_scores=0.39084595, pem=-1.5697736,
similarity=24.290009]
5: 187246 Santos_FC
    [pem_pos=0, final_scores=0.4537689,
final_local_scores=0.4139675, pem=-0.81493,
similarity=20.875046]
6: 6189919 Narciso_dos_Santos
    [pem_pos=1, final_scores=0.4623589,
final_local_scores=0.55450875, pem=-0.93916667,
similarity=25.744694]
7: 390875 São_Paulo
    [pem_pos=0, final_scores=0.33531654,
final_local_scores=0.29958573, pem=0.0, similarity=13.029948]
8: 25936 Rio_de_Janeiro
    [pem_pos=0, final_scores=0.4362145,
final_local_scores=0.40671265, pem=-0.1260924,
similarity=16.912539]
9: 681926 Brazilian_Football_Confederation
    [pem_pos=0, final_scores=0.37496084,
final_local_scores=0.38346505, pem=0.0, similarity=15.533074]
10: 681926 Brazilian_Football_Confederation
    [pem_pos=0, final_scores=0.40971518,
final_local_scores=0.42142886, pem=-0.3957313,
similarity=18.818039]
11: 25936 Rio_de_Janeiro
    [pem_pos=0, final_scores=0.48942906,
```

```
final_local_scores=0.46484149, pem=-1.4807417,
similarity=26.014013]
12: 19004 Moscow
    [pem_pos=0, final_scores=0.6183445,
final_local_scores=0.69901365, pem=-0.106582455,
similarity=25.529287]
13: 149286 Brazil_national_football_team
    [pem_pos=1, final_scores=0.47249413,
final_local_scores=0.43915206, pem=-2.5928838,
similarity=31.295393]
14: 25127146 Russia_national_football_team
    [pem_pos=2, final_scores=0.52084714,
final_local_scores=0.5566968, pem=-3.5373242,
similarity=39.939175]
15: 187246 Santos_FC
    [pem_pos=0, final_scores=0.46632165,
final_local_scores=0.42767948, pem=-0.81493,
similarity=21.284239]
16: 6189919 Narciso_dos_Santos
    [pem_pos=1, final_scores=0.41918707,
final_local_scores=0.5073501, pem=-0.93916667,
similarity=24.337385]
17: 187274 São_Paulo_FC
    [pem_pos=2, final_scores=0.47468597,
final_local_scores=0.43336207, pem=-1.5697736,
similarity=25.558773]
18: 187274 São_Paulo_FC
    [pem_pos=2, final_scores=0.49949354,
final_local_scores=0.46046048, pem=-1.5697736,
similarity=26.367443]
19: 1577507 Esporte_Clube_Juventude
    [pem_pos=0, final_scores=0.45332378,
final_local_scores=0.4545976, pem=0.0, similarity=17.655811]
20: 187291 Sport_Club_Internacional
    [pem_pos=0, final_scores=0.4273184,
final_local_scores=0.39461148, pem=0.0, similarity=15.865706]
21: 187241 Sport_Club_Corinthians_Paulista
    [pem_pos=0, final_scores=0.37753397,
final_local_scores=0.32503915, pem=-0.017981783,
similarity=13.887318]
22: 26667 Spain
    [pem_pos=0, final_scores=0.55531734,
final_local_scores=0.5820783, pem=-0.18857051,
similarity=22.485565]
23: 26667 Spain
    [pem_pos=0, final_scores=0.58368415,
final_local_scores=0.61306465, pem=-0.18857051,
similarity=23.410261]
24: 390875 São_Paulo
    [pem_pos=0, final_scores=0.40128034,
final_local_scores=0.37164122, pem=0.0, similarity=15.180228]
25: 439002 Curitiba
```

```
[pem_pos=0, final_scores=0.46648616,
final_local_scores=0.45632565, pem=0.0, similarity=17.70738]
26: 187346 Clube_Atlético_Paranaense
    [pem_pos=0, final_scores=0.4062739,
final_local_scores=0.38067347, pem=0.0, similarity=15.449768]
27: 3383 Brazil
    [pem_pos=0, final_scores=0.67105967,
final_local_scores=0.6920119, pem=-0.2733083,
similarity=26.22702]
28: 187238 Botafogo_de_Futebol_e_Regatas
    [pem_pos=0, final_scores=0.31268495,
final_local_scores=0.2649757, pem=-0.059020996,
similarity=12.318081]
29: 2335808 Túlio_Maravilha
    [pem_pos=2, final_scores=0.5225852,
final_local_scores=0.5567747, pem=-1.6335354,
similarity=29.5884]
30: 9647657 Netherlands_national_football_team
    [pem_pos=1, final_scores=0.5200934,
final_local_scores=0.5555999, pem=-2.660505,
similarity=35.138165]
31: 187328 Esporte_Clube_Bahia
    [pem_pos=1, final_scores=0.29253882,
final_local_scores=0.24583712, pem=-1.7376727,
similarity=20.875717]
32: 2335808 Túlio_Maravilha
    [pem_pos=2, final_scores=0.51589304,
final_local_scores=0.5494645, pem=-1.6335354,
similarity=29.37025]
```

Sample: b'427_SOCCER'

```
SOCCER - [1 1 ARGENTINE 1] CHAMPIONSHIP RESULTS . [2 2 BUENOS
AIRES 2] 1996-08-26 Results of matches on the opening weekend
of the Argentine Apertura championship : [3 3 Estudiantes 3]
2 [4 4 Boca Juniors 4] 3 [5 5 Ferro Carril Oeste 5] 0 [6 6 In
dependiente 6] 3 Gimnasia-Jujuy 1 [7 7 Platense 7] 0 [8 8 Hur
acan 8] 0 [9 9 Lanus 9] 0 Huracan-Corrientes 3 1 Union 6
Newell 's Old Boys 0 [10 10 Velez Sarsfield 10] 2 [11 11 Raci
ng Club 11] 0 [12 12 Rosario Central 12] 2 [13 13 River Plate
13] 0 Gimnasia-La Plata 0 [14 14 San Lorenzo 14] 0 [15 15 Ban
field 15] 1 Playing Monday : [16 16 Deportivo Espanol 16] v
Colon Note : the Apertura is the first of two championships
played in the [17 17 Argentine 17] season . The teams meet
each other once in each tournament . There is no overall
champion .
```

```
1: 18951905 Argentina
2: 255919 Buenos_Aires
3: 1081561 Estudiantes_de_La_Plata
4: 401877 Boca_Juniors
5: 2149745 Club_Ferro_Carril_Oeste
6: 1121856 Club_Atlético_Independiente
7: 3954167 Club_Atlético_Platense
```

```
8: 2433314 Club_Atlético_Huracán
9: 1133768 Club_Atlético_Lanús
10: 1482035 Club_Atlético_Vélez_Sarsfield
11: 1404901 Racing_Club_de_Avellaneda
12: 1100763 Rosario_Central
13: 740900 Club_Atlético_River_Plate
14: 783031 San_Lorenzo_de_Almagro
15: 1822348 Club_Atlético_Banfield
16: 7213236 Deportivo_Español
17: 18951905 Argentina
1: 1255902 Unión_de_Santa_Fe
    [winner: 481130 Union_(American_Civil_War), pem_pos=0,
final_scores=-0.81440365, final_local_scores=-0.75032276,
pem=-1.132797, similarity=-12.141089]
    RECALL MISS: 1255902 Unión_de_Santa_Fe]
1: 18951905 Argentina
    [pem_pos=0, final_scores=0.73506826,
final_local_scores=0.7798118, pem=-0.18885723,
similarity=28.387888]
2: 255919 Buenos_Aires
    [pem_pos=0, final_scores=0.41678983,
final_local_scores=0.421448, pem=-0.12597719,
similarity=17.351645]
3: 1081561 Estudiantes_de_La_Plata
    [pem_pos=0, final_scores=0.36441153,
final_local_scores=0.30062425, pem=-0.054989595,
similarity=13.3599825]
4: 401877 Boca_Juniors
    [pem_pos=0, final_scores=0.39461106,
final_local_scores=0.32838342, pem=0.0, similarity=13.889329]
5: 2149745 Club_Ferro_Carril_Oeste
    [pem_pos=0, final_scores=0.41793972,
final_local_scores=0.3811304, pem=0.0, similarity=15.463404]
6: 1121856 Club_Atlético_Independiente
    [pem_pos=0, final_scores=0.42652488,
final_local_scores=0.3630914, pem=0.0, similarity=14.925085]
7: 3954167 Club_Atlético_Platense
    [pem_pos=0, final_scores=0.23813729,
final_local_scores=0.19659084, pem=-0.32154775,
similarity=11.705002]
8: 2433314 Club_Atlético_Huracán
    [pem_pos=1, final_scores=0.43148297,
final_local_scores=0.376935, pem=-0.56511337,
similarity=18.411379]
9: 1133768 Club_Atlético_Lanús
    [pem_pos=1, final_scores=0.5235551,
final_local_scores=0.47894967, pem=-0.052490193,
similarity=18.667976]
10: 1482035 Club_Atlético_Vélez_Sarsfield
    [pem_pos=0, final_scores=0.5630079,
final_local_scores=0.5181303, pem=0.0, similarity=19.551754]
11: 1404901 Racing_Club_de_Avellaneda
```

```
[pem_pos=0, final_scores=0.43348354,
final_local_scores=0.37319088, pem=0.0, similarity=15.226473]
12: 1100763 Rosario_Central
    [pem_pos=0, final_scores=0.46198624,
final_local_scores=0.40759057, pem=0.0, similarity=16.253029]
13: 740900 Club_Atlético_River_Plate
    [pem_pos=0, final_scores=0.48350322,
final_local_scores=0.42736566, pem=-0.038094174,
similarity=17.05032]
14: 783031 San_Lorenzo_de_Almagro
    [pem_pos=0, final_scores=0.4670815,
final_local_scores=0.40644473, pem=-0.71188563,
similarity=20.09018]
15: 1822348 Club_Atlético_Banfield
    [pem_pos=0, final_scores=0.4391287,
final_local_scores=0.3861571, pem=-0.29415068,
similarity=17.213047]
16: 7213236 Deportivo_Español
    [pem_pos=0, final_scores=0.3104692,
final_local_scores=0.28978628, pem=0.0, similarity=12.737514]
17: 18951905 Argentina
    [pem_pos=0, final_scores=0.42158931,
final_local_scores=0.4373837, pem=-0.18885723,
similarity=18.169151]
```

Sample: b'428_SOCCER'

```
SOCCER - [1 1 HONDURAS 1] BEAT [2 2 CUBA 2] 4-0 IN FRIENDLY . [3 3 TEGUCIGALPA 3] 1996-08-26 [4 4 Honduras 4] beat [5 5 Cuba 5] 4-0 ( halftime 3-0 ) in a friendly soccer international on Sunday . Scorers : Juan Castro ( 3rd minute ) , Enrique Centeno ( 33rd and 84th ) , [6 6 Carlos Pavón 6] ( 37th )
1: 13394 Honduras
2: 5042481 Cuba
3: 25994100 Tegucigalpa
4: 1013983 Honduras_national_football_team
5: 1024940 Cuba_national_football_team
6: 1125764 Carlos_Pavón
1: 13394 Honduras
    [pem_pos=0, final_scores=0.51450795,
final_local_scores=0.5093622, pem=-0.30668682,
similarity=20.957909]
2: 5042481 Cuba
    [pem_pos=0, final_scores=0.4282896,
final_local_scores=0.4372763, pem=-0.23907764,
similarity=18.439053]
3: 25994100 Tegucigalpa
    [pem_pos=0, final_scores=0.39780647,
final_local_scores=0.41042894, pem=0.0, similarity=16.33773]
4: 1013983 Honduras_national_football_team
    [pem_pos=1, final_scores=0.42065334,
final_local_scores=0.42200547, pem=-1.8252479,
```

```
similarity=26.60918]
5: 1024940 Cuba_national_football_team
    [pem_pos=1, final_scores=0.5268528,
final_local_scores=0.54896265, pem=-3.3116636,
similarity=38.481194]
6: 1125764 Carlos_Pavón
    [pem_pos=0, final_scores=0.43311512,
final_local_scores=0.47575217, pem=0.0, similarity=18.287107]
```

Sample: b'429_SOCCER'

SOCCER - [1 1 BRAZILIAN 1] CHAMPIONSHIP RESULTS /
STANDINGS . [2 2 RIO DE JANEIRO 2] 1996-08-26 Results of [3 3
Brazilian 3] soccer championship matches played at the
weekend . [4 4 Bahia 4] 1 [5 5 Botafogo 5] 2 [6 6 Bragantino 6
] 1 [7 7 Vasco da Gama 7] 2 Cricuma 4 [8 8 Fluminense 8] 1 [9
9 Cruzeiro 9] 2 [10 10 Flamengo 10] 1 [11 11 Goias 11] 0 [12 1
2 Palmeiras 12] 0 [13 13 Gremio 13] 2 [14 14 Vitoria 14] 2 [15
15 Juventude 15] 2 [16 16 Internacional 16] 1 [17 17 Parana 17
] 3 [18 18 Guarani 18] 0 [19 19 Portuguesa 19] 3 [20 20 Atlet
ico Mineiro 20] 1 [21 21 Sao Paulo 21] 2 [22 22 Santos 22] 1 [
23 23 Sport Recife 23] 3 [24 24 Coritiba 24] 0 Standings
(tabulated under played , won , drawn , lost , goals for ,
against , points) : [25 25 Sao Paulo 25] 4 3 1 0 10 5 10 [26
26 Juventude 26] 5 3 1 1 5 4 10 [27 27 Portuguesa 27] 4 3 0 1
8 3 9 [28 28 Palmeiras 28] 5 2 3 0 8 1 9 [29 29 Goias 29] 5 2
2 1 7 4 8 [30 30 Gremio 30] 3 2 1 0 11 4 7 [31 31 Cruzeiro 31]
3 2 1 0 4 2 7 [32 32 Sport Recife 32] 5 2 1 2 7 6 7 [33 33 Pa
rana 33] 4 2 1 2 5 5 7 [34 34 Flamengo 34] 4 2 0 2 4 4 6 [35 3
5 Atletico Mineiro 35] 5 2 0 3 6 7 6 [36 36 Vasco da Gama 36]
4 2 0 2 6 7 6 [37 37 Coritiba 37] 5 2 0 3 3 9 6 [38 38 Botafo
go 38] 3 1 2 0 4 3 5 [39 39 Internacional 39] 4 1 2 1 4 3 5 [4
0 40 Criciuma 40] 5 1 2 2 7 8 5 [41 41 Vitoria 41] 5 1 2 2 5
6 5 [42 42 Santos 42] 3 1 1 1 3 3 4 [43 43 Corinthians 43] 4
1 1 2 1 3 4 [44 44 Bahia 44] 5 1 1 3 5 8 4 [45 45 Fluminense 4
5] 4 1 1 2 3 6 4 [46 46 Atletico Paranaense 46] 3 1 0 2 4 6 3
[47 47 Guarani 47] 3 0 1 2 1 5 1 [48 48 Bragantino 48] 4 0 0
4 3 12 0 Note : Top eight teams qualify for the quarter-
finals . If teams are level on points , positions are
determined by the number of games won .

```
1: 3383 Brazil
2: 25936 Rio_de_Janeiro
3: 3383 Brazil
4: 187328 Esporte_Clube_Bahia
5: 187238 Botafogo_de_Futebol_e_Regatas
6: 1861824 Clube_Atlético_Bragantino
7: 187324 CR_Vasco_da_Gama
8: 187313 Fluminense_FC
9: 187301 Cruzeiro_Esporte_Clube
10: 187239 Clube_de_Regatas_do_Flamengo
11: 1577470 Goiás_Esporte_Clube
12: 187243 Sociedade_Esportiva_Palmeiras
13: 18564144 Grêmio_Foot-Ball_Porto_Alegrense
```

```
14: 187310 Esporte_Clube_Vitória
15: 1577507 Esporte_Clube_Juventude
16: 187291 Sport_Club_Internacional
17: 311141 Paraná_Clube
18: 1367196 Guarani_Futebol_Clube
19: 1742904 Associação_Portuguesa_de_Desportos
20: 187305 Clube_Atlético_Mineiro
21: 187274 São_Paulo_FC
22: 187246 Santos_FC
23: 604575 Sport_Club_do_Recife
24: 1550341 Coritiba_Foot_Ball_Club
25: 187274 São_Paulo_FC
26: 1577507 Esporte_Clube_Juventude
27: 1742904 Associação_Portuguesa_de_Desportos
28: 187243 Sociedade_Esportiva_Palmeiras
29: 1577470 Goiás_Esporte_Clube
30: 18564144 Grêmio_Foot-Ball_Porto_Alegrense
31: 187301 Cruzeiro_Esporte_Clube
32: 604575 Sport_Club_do_Recife
33: 311141 Paraná_Clube
34: 187239 Clube_de_Regatas_do_Flamengo
35: 187305 Clube_Atlético_Mineiro
36: 187324 CR_Vasco_da_Gama
37: 1550341 Coritiba_Foot_Ball_Club
38: 187238 Botafogo_de_Futebol_e_Regatas
39: 187291 Sport_Club_Internacional
40: 1737058 Criciúma_Esporte_Clube
41: 187310 Esporte_Clube_Vitória
42: 187246 Santos_FC
43: 187241 Sport_Club_Corinthians_Paulista
44: 187328 Esporte_Clube_Bahia
45: 187313 Fluminense_FC
46: 187346 Clube_Atlético_Paranaense
47: 1367196 Guarani_Futebol_Clube
48: 1861824 Clube_Atlético_Bragantino
1: 3383 Brazil
    [pem_pos=0, final_scores=0.76675755,
final_local_scores=0.85001034, pem=-0.2733083,
similarity=30.942009]
2: 25936 Rio_de_Janeiro
    [pem_pos=0, final_scores=0.45953554,
final_local_scores=0.49044335, pem=0.0, similarity=18.725521]
3: 3383 Brazil
    [pem_pos=0, final_scores=0.5844828,
final_local_scores=0.65090287, pem=-0.2733083,
similarity=25.000246]
4: 187328 Esporte_Clube_Bahia
    [pem_pos=1, final_scores=0.3288414,
final_local_scores=0.25647286, pem=-1.7376727,
similarity=21.19311]
5: 187238 Botafogo_de_Futebol_e_Regatas
    [pem_pos=0, final_scores=0.38147336,
```

```
final_local_scores=0.3123994, pem=-0.059020996,
similarity=13.733298]
6: 1861824 Clube_Atlético_Bragantino
    [pem_pos=0, final_scores=0.37786323,
final_local_scores=0.3536903, pem=0.0, similarity=14.644537]
7: 187324 CR_Vasco_da_Gama
    [pem_pos=1, final_scores=0.30248028,
final_local_scores=0.22692361, pem=-0.725215,
similarity=14.805398]
8: 187313 Fluminense_FC
    [pem_pos=0, final_scores=0.38760442,
final_local_scores=0.31637987, pem=0.0, similarity=13.531119]
9: 187301 Cruzeiro_Esporte_Clube
    [pem_pos=0, final_scores=0.46553594,
final_local_scores=0.40513313, pem=0.0, similarity=16.179693]
10: 187239 Clube_de_Regatas_do_Flamengo
    [pem_pos=0, final_scores=0.39133126,
final_local_scores=0.3245287, pem=0.0, similarity=13.774297]
11: 1577470 Goiás_Esporte_Clube
    [pem_pos=1, final_scores=0.4644428,
final_local_scores=0.42548615, pem=-0.957059,
similarity=21.991703]
12: 187243 Sociedade_Esportiva_Palmeiras
    [pem_pos=0, final_scores=0.45076406,
final_local_scores=0.38819462, pem=0.0, similarity=15.674214]
13: 18564144 Grêmio_Foot-Ball_Porto_Alegrense
    [pem_pos=0, final_scores=0.52741426,
final_local_scores=0.47199655, pem=0.0, similarity=18.175032]
14: 187310 Esporte_Clube_Vitória
    [pem_pos=2, final_scores=0.34960514,
final_local_scores=0.2831039, pem=-1.938656,
similarity=23.08081]
15: 1577507 Esporte_Clube_Juventude
    [pem_pos=0, final_scores=0.46942323,
final_local_scores=0.43237633, pem=0.0, similarity=16.992685]
16: 187291 Sport_Club_Internacional
    [pem_pos=0, final_scores=0.44869298,
final_local_scores=0.38742733, pem=0.0, similarity=15.651318]
17: 311141 Paraná_Clube
    [pem_pos=4, final_scores=0.515928,
final_local_scores=0.48934, pem=-1.6926686,
similarity=27.89759]
18: 1367196 Guarani_Futebol_Clube
    [pem_pos=0, final_scores=0.36335325,
final_local_scores=0.30504325, pem=-0.25016624,
similarity=14.553255]
19: 1742904 Associação_Portuguesa_de_Desportos
    [pem_pos=0, final_scores=0.34856647,
final_local_scores=0.28833517, pem=-0.14295863,
similarity=13.471641]
20: 187305 Clube_Atlético_Mineiro
    [pem_pos=0, final_scores=0.36817586,
```

```
final_local_scores=0.29854572, pem=0.0, similarity=12.998913]
21: 187274 São_Paulo_FC
    [pem_pos=2, final_scores=0.48843098,
final_local_scores=0.43083555, pem=-1.5697736,
similarity=25.483376]
22: 187246 Santos_FC
    [pem_pos=0, final_scores=0.36561382,
final_local_scores=0.30576032, pem=-0.81493,
similarity=17.645927]
23: 604575 Sport_Club_do_Recife
    [pem_pos=0, final_scores=0.4949836,
final_local_scores=0.44652593, pem=0.0, similarity=17.414938]
24: 1550341 Coritiba_Foot_Ball_Club
    [pem_pos=0, final_scores=0.54280657,
final_local_scores=0.49961424, pem=0.0, similarity=18.999199]
25: 187274 São_Paulo_FC
    [pem_pos=2, final_scores=0.4593398,
final_local_scores=0.3990578, pem=-1.5697736,
similarity=24.535067]
26: 1577507 Esporte_Clube_Juventude
    [pem_pos=0, final_scores=0.45743853,
final_local_scores=0.41928488, pem=0.0, similarity=16.60201]
27: 1742904 Associação_Portuguesa_de_Desportos
    [pem_pos=0, final_scores=0.35993916,
final_local_scores=0.30075806, pem=-0.14295863,
similarity=13.842364]
28: 187243 Sociedade_Esportiva_Palmeiras
    [pem_pos=0, final_scores=0.43301243,
final_local_scores=0.36880362, pem=0.0, similarity=15.095549]
29: 1577470 Goiás_Esporte_Clube
    [pem_pos=1, final_scores=0.41048682,
final_local_scores=0.3665474, pem=-0.957059,
similarity=20.232853]
30: 18564144 Grêmio_Foot-Ball_Porto_Alegrense
    [pem_pos=0, final_scores=0.5109485,
final_local_scores=0.4540102, pem=0.0, similarity=17.638283]
31: 187301 Cruzeiro_Esporte_Clube
    [pem_pos=0, final_scores=0.4449417,
final_local_scores=0.38263702, pem=0.0, similarity=15.508365]
32: 604575 Sport_Club_do_Recife
    [pem_pos=0, final_scores=0.49078882,
final_local_scores=0.4419437, pem=0.0, similarity=17.278194]
33: 311141 Paraná_Clube
    [pem_pos=4, final_scores=0.49198675,
final_local_scores=0.46318787, pem=-1.6926686,
similarity=27.117157]
34: 187239 Clube_de_Regatas_do_Flamengo
    [pem_pos=0, final_scores=0.39468163,
final_local_scores=0.3281884, pem=0.0, similarity=13.883509]
35: 187305 Clube_Atlético_Mineiro
    [pem_pos=0, final_scores=0.35057044,
final_local_scores=0.2793145, pem=0.0, similarity=12.4250145]
```

```
36: 187324 CR_Vasco_da_Gama
    [pem_pos=1, final_scores=0.30409354,
final_local_scores=0.22868592, pem=-0.725215,
similarity=14.857988]
37: 1550341 Coritiba_Foot_Ball_Club
    [pem_pos=0, final_scores=0.52314156,
final_local_scores=0.4781332, pem=0.0, similarity=18.358162]
38: 187238 Botafogo_de_Futebol_e_Regatas
    [pem_pos=0, final_scores=0.38427013,
final_local_scores=0.31545442, pem=-0.059020996,
similarity=13.824467]
39: 187291 Sport_Club_Internacional
    [pem_pos=0, final_scores=0.46152443,
final_local_scores=0.40144372, pem=0.0, similarity=16.069593]
40: 1737058 Criciúma_Esporte_Clube
    [pem_pos=1, final_scores=0.5056971,
final_local_scores=0.4742722, pem=-0.523729,
similarity=21.091063]
41: 187310 Esporte_Clube_Vitória
    [pem_pos=2, final_scores=0.34549356,
final_local_scores=0.27861258, pem=-1.938656,
similarity=22.946781]
42: 187246 Santos_FC
    [pem_pos=0, final_scores=0.3716573,
final_local_scores=0.3123619, pem=-0.81493,
similarity=17.842932]
43: 187241 Sport_Club_Corinthians_Paulista
    [pem_pos=0, final_scores=0.44062805,
final_local_scores=0.37566882, pem=-0.017981783,
similarity=15.398207]
44: 187328 Esporte_Clube_Bahia
    [pem_pos=1, final_scores=0.28003293,
final_local_scores=0.20315698, pem=-1.7376727,
similarity=19.602057]
45: 187313 Fluminense_FC
    [pem_pos=0, final_scores=0.41653925,
final_local_scores=0.34798676, pem=0.0, similarity=14.474332]
46: 187346 Clube_Atlético_Paranaense
    [pem_pos=0, final_scores=0.40353364,
final_local_scores=0.34492576, pem=0.0, similarity=14.382986]
47: 1367196 Guarani_Futebol_Clube
    [pem_pos=0, final_scores=0.37935525,
final_local_scores=0.322523, pem=-0.25016624,
similarity=15.074884]
48: 1861824 Clube_Atlético_Bragantino
    [pem_pos=0, final_scores=0.34529907,
final_local_scores=0.31811893, pem=0.0, similarity=13.583016]
```

Sample: b'430_BASKETBALL'

BASKETBALL - [1 1 PHILIPPINE 1] PRO-LEAGUE RESULTS . MANILA
1996-08-26 Results of semi-final round games played on late
Sunday in the [2 2 Philippine Basketball Association 2]

```

second conference , which includes [3 3 American 3] players :
[4 4 Formula Shell 4] beat [5 5 Ginebra San Miguel 5] 89-86
( 45-46 half-time )
1: 23440 Philippines
2: 364158 Philippine_Basketball_Association
3: 3434750 United_States
4: 4092880 Shell_Turbo_Chargers
5: 3647082 Barangay_Ginebra_San_Miguel
1: 23440 Philippines
    [pem_pos=0, final_scores=0.48345727,
final_local_scores=0.5520968, pem=-0.5672473,
similarity=23.650162]
2: 364158 Philippine_Basketball_Association
    [pem_pos=0, final_scores=0.34541255,
final_local_scores=0.33631688, pem=0.0, similarity=14.12608]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.3953625,
final_local_scores=0.49329865, pem=-0.12994486,
similarity=19.51739]
4: 4092880 Shell_Turbo_Chargers
    [pem_pos=0, final_scores=0.5650632,
final_local_scores=0.59939134, pem=0.0, similarity=21.976746]
5: 3647082 Barangay_Ginebra_San_Miguel
    [pem_pos=0, final_scores=0.35915297,
final_local_scores=0.35757977, pem=0.0, similarity=14.760607]

```

Sample: b'431_SOCCER'

SOCCER - RESULTS OF [1 1 SOUTH KOREAN 1] PRO-SOCCER GAMES . [
2 2 SEOUL 2] 1996-08-26 Results of [3 3 South Korean 3] pro-
soccer games played on Sunday . Puchon 3 Chonan 0 (halftime
1-0) [4 4 Pohang 4] 3 Chonbuk 2 (halftime 0-0) Standings
after games played on Sunday (tabulate under - won , drawn ,
lost , goals for , goals against , points) : W D L G / F G /
A P Puchon 2 1 0 4 0 7 Chonan 2 0 1 9 9 6 [5 5 Pohang 5] 1 1
1 8 8 4 [6 6 Ulsan 6] 1 0 1 6 6 3 Anyang 0 3 0 5 5 3 [7 7 Suw
on 7] 0 3 0 3 3 3 Pusan 0 2 0 3 3 2 Chonnam 0 2 1 4 5 2
Chonbuk 0 0 2 2 5 0

```

1: 27019 South_Korea
2: 19159283 Seoul
3: 27019 South_Korea
4: 1056366 Pohang_Steelers
5: 1056366 Pohang_Steelers
6: 1056425 Ulsan_Hyundai_FC
7: 1054817 Suwon_Samsung_Bluewings
1: 27019 South_Korea
    [pem_pos=0, final_scores=0.46566278,
final_local_scores=0.45768613, pem=0.0, similarity=17.74798]
2: 19159283 Seoul
    [pem_pos=0, final_scores=0.47687203,
final_local_scores=0.49140477, pem=-0.0378014,
similarity=18.959782]
3: 27019 South_Korea

```

```
[pem_pos=0, final_scores=0.513115,
final_local_scores=0.5095205, pem=0.0, similarity=19.29482]
4: 1056366 Pohang_Steelers
    [pem_pos=2, final_scores=0.66504544,
final_local_scores=0.66280854, pem=-1.0587178,
similarity=29.626715]
5: 1056366 Pohang_Steelers
    [pem_pos=2, final_scores=0.64924675,
final_local_scores=0.6455508, pem=-1.0587178,
similarity=29.111708]
6: 1056425 Ulsan_Hyundai_FC
    [pem_pos=3, final_scores=0.5600098,
final_local_scores=0.5516384, pem=-1.7336445,
similarity=29.979534]
7: 1054817 Suwon_Samsung_Bluewings
    [pem_pos=2, final_scores=0.4843034,
final_local_scores=0.46850747, pem=-1.0623775,
similarity=23.848284]
```

Sample: b'432_BASEBALL'

BASEBALL - RESULTS OF [1 1 S. KOREAN 1] PRO-BASEBALL GAMES .
[2 2 SEOUL 2] 1996-08-26 Results of [3 3 South Korean 3] pro-baseball games played on Sunday . 0B 2 Lotte 1 [4 4 Hanwha 4]
3 Haitai 2 Hyundai 8 Samsung 1 Ssangbangwool 3 LG 1 Standings
after games played on Sunday (tabulate under won , drawn ,
lost , winning percentage , games behind first place) W D L
PCT GB Haitai 63 2 41 .604 - Ssangbangwool 58 2 47 .551 5 1/2
Hyundai 56 5 47 .542 6 1/2 [5 5 Hanwha 5] 56 1 48 .538 7
Samsung 47 5 55 .463 15 Lotte 44 6 53 .456 15 1/2 LG 44 5 58 .
435 18 0B 41 6 60 .411 20 1/2

```
1: 27019 South_Korea
2: 19159283 Seoul
3: 27019 South_Korea
4: 4086558 Hanwha_Eagles
5: 4086558 Hanwha_Eagles
1: 27019 South_Korea
    [pem_pos=0, final_scores=0.5700926,
final_local_scores=0.55340225, pem=0.0, similarity=20.60434]
2: 19159283 Seoul
    [pem_pos=0, final_scores=0.49555224,
final_local_scores=0.49118984, pem=-0.0378014,
similarity=18.953367]
3: 27019 South_Korea
    [pem_pos=0, final_scores=0.52925897,
final_local_scores=0.5087977, pem=0.0, similarity=19.27325]
4: 4086558 Hanwha_Eagles
    [pem_pos=1, final_scores=0.59033716,
final_local_scores=0.62479115, pem=-0.6931472,
similarity=26.50417]
5: 4086558 Hanwha_Eagles
    [pem_pos=1, final_scores=0.5451195,
final_local_scores=0.5753978, pem=-0.6931472,
```

similarity=25.030172]

Sample: b'433_SOCCER'
SOCCER - [1 1 MOROCCAN 1] FIRST DIVISION RESULTS . [2 2 RABAT
2] 1996-08-26 Results of [3 3 Moroccan 3] first division
soccer matches played on Sunday : Widad Fes 3 Oujda 1 [4 4 Ra
ja Casablanca 4] 4 Tetouan 0 [5 5 Jeunesse Massira 5] 0 [6 6 W
idad Casablanca 6] 2 Sporting Sale 0 Meknes 1 Settat 1
Marrakesh 0 [7 7 Khouribga 7] 3 Mohammedia 0 Sidi Kacem 0
Royal Armed Forces 0 El Jadida 1 [8 8 Hassania Agadir 8] 0
1: 19291 Morocco
2: 57400 Rabat
3: 19291 Morocco
4: 1689806 Raja_Casablanca
5: 6230254 JS_Massira
6: 3884066 Wydad_Casablanca
7: 6230182 OC_Khouribga
8: 4930399 Hassania_Agadir
1: 19291 Morocco
[pem_pos=0, final_scores=0.59125215,
final_local_scores=0.61732167, pem=-0.2674848,
similarity=23.966448]
2: 57400 Rabat
[pem_pos=0, final_scores=0.38732696,
final_local_scores=0.3849125, pem=-0.077565834,
similarity=15.998085]
3: 19291 Morocco
[pem_pos=0, final_scores=0.6007518,
final_local_scores=0.6276986, pem=-0.2674848,
similarity=24.276115]
4: 1689806 Raja_Casablanca
[pem_pos=0, final_scores=0.3977849,
final_local_scores=0.37907702, pem=0.0, similarity=15.402127]
5: 6230254 JS_Massira
[pem_pos=0, final_scores=0.30379993,
final_local_scores=0.34572, pem=0.0, similarity=14.406687]
6: 3884066 Wydad_Casablanca
[pem_pos=0, final_scores=0.5290461,
final_local_scores=0.53828293, pem=0.0, similarity=20.153149]
7: 6230182 OC_Khouribga
[pem_pos=1, final_scores=0.33041197,
final_local_scores=0.35708034, pem=-1.0786515,
similarity=20.611578]
8: 4930399 Hassania_Agadir
[pem_pos=0, final_scores=0.4745828,
final_local_scores=0.5032717, pem=0.0, similarity=19.108345]

Sample: b'434_TENNIS'

TENNIS - QUENCH YOUR THIRST - IF YOU CAN AFFORD IT AT [1 1 U.
S. OPEN 1] . Bill Berkrot [2 2 NEW YORK 2] 1996-08-26 A
message on television monitors all around the National Tennis
Centre reads : " Due to hot weather please seek shade and

drink plenty of fluids " -- sound advice until you check out the price of fluids . Perhaps the advisory was cut off before concluding : " ...and bring plenty of money . " A small bottle of a garishly-coloured sports drink at the sun-drenched [3 3 U.S. Open 3] is going for \$ 3.75 , while a litre of basic , life-sustaining water will set you back \$ 4.00 -- for water ? " At the 1 Olympics water was only a dollar , and that was the [4 4 Olympics 4] , " said one incredulous fan , noting that the [5 5 Atlanta Games 5] had been notorious for price gouging . [6 6 U.S. Open 6] officials managed to insult most of the male tennis players last week with their controversial handling of the seeding and draw . When the tournament began on Monday it was the fans ' turn to be offended . " That baked lasagna better be good for \$ 8.50 , " said [7 7 New Yorker 7] Rebecca Weinstein , a [8 8 U.S. Open 8] regular who was eating a sandwich she had brought from home . A trio of hungry fans at the food court who had already forked over the lasagna money pronounced it good , but Carol Perry chimed in , " The water is ridiculous , they want four dollars for the water , you might as well get a glass of wine . " Indeed , a nice glass of chardonnay or white zinfandel was going for \$ 4.75 , while an imported beer was just a bit more than the water at \$ 4.50 . What 's the message here ? " Maybe they want us to be alcoholics , " Perry joked before lifting her glass of wine . Fans will be shelling out \$ 12.50 for a hamburger and a large french fries . And that little snack is guaranteed to make you thirsty . Make that \$ 16.50 . Even a sandwich as pedestrian as a ham and swiss cheese is going for a whopping \$ 8.00 . Of course , it is served on a tuscan roll . It must be the cost of flying those rolls over from [9 9 Tuscany 9] every day that drives up the price of the sandwich .

```
1: 82729 US_Open_(tennis)
2: 645042 New_York_City
3: 82729 US_Open_(tennis)
4: 27566 Summer_Olympic_Games
5: 72311 1996_Summer_Olympics
6: 82729 US_Open_(tennis)
7: 645042 New_York_City
8: 82729 US_Open_(tennis)
9: 21967242 Tuscany
1: 72311 1996_Summer_Olympics
    [winner: 72311 1996_Summer_Olympics, pem_pos=14,
final_scores=0.008811042, final_local_scores=-0.033421993,
pem=-4.001833, similarity=24.854923]
    gt: 72311 1996_Summer_Olympics, pem_pos=14,
final_scores=0.008811042, final_local_scores=-0.033421993,
pem=-4.001833, similarity=24.854923 ]
1: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.5517179,
final_local_scores=0.52483094, pem=-1.2370842,
similarity=26.47917]
```

```
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.4338128,
final_local_scores=0.45685512, pem=-1.7906687,
similarity=27.461115]
3: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.51913625,
final_local_scores=0.48924047, pem=-1.2370842,
similarity=25.41708]
4: 27566 Summer_Olympic_Games
    [pem_pos=2, final_scores=0.2763238,
final_local_scores=0.27320915, pem=-2.7333555,
similarity=27.107231]
5: 72311 1996_Summer_Olympics
    [pem_pos=0, final_scores=0.4018445,
final_local_scores=0.42036372, pem=0.0, similarity=16.634205]
6: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.54129785,
final_local_scores=0.51344866, pem=-1.2370842,
similarity=26.1395]
7: 645042 New_York_City
    [pem_pos=1, final_scores=0.35851008,
final_local_scores=0.37459838, pem=-1.1225154,
similarity=21.37289]
8: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.49689394,
final_local_scores=0.46494412, pem=-1.2370842,
similarity=24.692028]
9: 21967242 Tuscany
    [pem_pos=0, final_scores=0.38978636,
final_local_scores=0.51377267, pem=-0.024571078,
similarity=19.555336]
```

Sample: b'435_TENNIS'

TENNIS - [1 1 HUBER 1] AND [2 2 MALEEEVA 2] FALL , UP-AND-COMERS ADVANCE AT OPEN . Larry Fine [3 3 NEW YORK 3] 1996-08-26 [4 4 Martina Hingis 4] led a youthful charge and [5 5 Australian Open 5] finalist [6 6 Anke Huber 6] and [7 7 Magdalena Maleeva 7] were fallen seeds on Monday in a hot , sunny opening to the [8 8 U.S. Open 8] tennis championships . The 15-year-old [9 9 Hingis 9] , seeded 16th , was honoured to play the first match of the season 's last [10 10 Grand Slam 10] on Stadium Court but happy to hurry off with a straight-sets victory over the 112th-ranked [11 11 Angeles Montolio 11] of [12 12 Spain 12] . " It was very hot and I didn't want to stay long on the court , " said a cheery [13 13 Hingis 13] , who had no worries in racing to a 6-1 6-0 victory against the overmatched [14 14 Spaniard 14] . Hoping for a longer engagement on the cement at 1 Flushing Meadows were the sixth-seeded [15 15 Huber 15] of [16 16 Germany 16] and 12th seed [17 17 Maleeva 17] of [18 18 Bulgaria 18] . [19 19 Huber 19] , who lost to [20 20 Monica Seles 20] in the [21 21 Australian Open 21] final , fell victim to an unfortunate draw

in bowing to dangerous floater [22 22 Amanda Coetzer 22] of [2 3 23 South Africa 23] . [24 24 Coetzer 24] , ranked 17th , avenged her defeat to [25 25 Huber 25] in the [26 26 Australian Open 26] semifinals by winning 6-1 2-6 6-2 . " I looked at it as not a first round match , just a great challenge for me , " said [27 27 Coetzer 27] , 24 . " I was really concentrating on keeping my own momentum and my own rhythm . " She is tough to play in that way because she plays very up and down . She played one great game and than a few errors . The challenge was just for me to keep playing my own game . " [28 28 Huber 28] , who reached the final a week ago at [29 29 Manhattan Beach 29] , could only mourn her luck of the draw . " I was n't happy when I saw the draw . She was the first non- seeded player , " said the 21-year-old [30 30 German 30] . " It 's always tough to play somebody like that in the first round in a [31 31 Grand Slam 31] . " I think I did n't play that bad today . It was maybe my best first round match in a [32 32 Grand Slam 32] I ever played . " Monday brought the best out in [33 33 U.S. Open 33] rookie [34 34 Aleksandra Olsza 34] of [35 35 Poland 35] , ranked 110th . The 18-year-old [36 36 Olsza 36] , last year 's [37 37 Wimbledon 37] junior champion , celebrated her debut in the main draw of the Open by removing [38 38 Maleeva 38] 6-4 6-2 . The curtain-raising victories by [39 39 Hingis 39] and [40 40 Olsza 40] provided a ringing endorsement for the newest wave of women 's players coming up from the junior ranks . The [41 41 Swiss 41] teenager , a twice [42 42 French Open 42] junior champion and a [43 43 Wimbledon 43] juniors winner , had already proven her main stage mettle by reaching the quarters at this year 's [44 44 Australian Open 44] . " I hope I can get into the last 16 , " said [45 45 Hingis 45] , seeded to face third seed [46 46 Arantxa Sanchez Vicario 46] in the fourth round . [47 47 Hingis 47] has been working hard on conditioning and has lost eight pounds (3.5 kilos) in advance of the championships . " There will be tough matches but I hope I can get there , " she said . " Then we 'll see if [48 48 Arantxa 48] will be there , too . " The fast-moving [49 49 Olsza 49] , 18 , was cool in her opening match . " I was n't scared when I heard that I was playing [50 50 Maleeva 50] , " said [51 51 Olsza 51] . " I know that if I want to play professional tennis I have to do my best to try to beat her and I ca n't be scared . " [52 52 Olsza 52] is undaunted by the level of competition in the pros . " In terms of tennis , I think the junior players are really good now . In a few years , it could change a lot among the top players . " Two big-serving women 's players made quick work of [53 53 Japanese 53] opponents . [54 54 Brenda Schultz-McCarthy 54] of the [55 55 Netherlands 55] , the 13th seed , was a 6-1 6-4 winner over [56 56 Japan 56] 's Nana Miyaga , while [57 57 Czech 57] veteran [58 58 Helena Sukova 58] prevailed over Yone Kamio 6-2 6-3 . [59 59 Austrian 59] [60 60 Barbara Paulus 60] , seeded 14th , also reached the second round with a 6-2 6-1 victory over Yi

Jing-Qian of [61 61 China 61] .
1: 1443478 Anke_Huber
2: 1111175 Magdalena_Maleeva
3: 645042 New_York_City
4: 19047 Martina_Hingis
5: 97851 Australian_Open
6: 1443478 Anke_Huber
7: 1111175 Magdalena_Maleeva
8: 82729 US_Open_(tennis)
9: 19047 Martina_Hingis
10: 197638 Grand_Slam_(tennis)
11: 12450241 Angeles_Montolio
12: 26667 Spain
13: 19047 Martina_Hingis
14: 963152 Spanish_people
15: 1443478 Anke_Huber
16: 11867 Germany
17: 1111175 Magdalena_Maleeva
18: 3415 Bulgaria
19: 1443478 Anke_Huber
20: 300468 Monica_Seles
21: 97851 Australian_Open
22: 1951841 Amanda_Coetzer
23: 17416221 South_Africa
24: 1951841 Amanda_Coetzer
25: 1443478 Anke_Huber
26: 97851 Australian_Open
27: 1951841 Amanda_Coetzer
28: 1443478 Anke_Huber
29: 107656 Manhattan_Beach,_California
30: 152735 Germans
31: 197638 Grand_Slam_(tennis)
32: 197638 Grand_Slam_(tennis)
33: 82729 US_Open_(tennis)
34: 22207361 Aleksandra_Olsza
35: 22936 Poland
36: 22207361 Aleksandra_Olsza
37: 91195 The_Championships,_Wimbledon
38: 1111175 Magdalena_Maleeva
39: 19047 Martina_Hingis
40: 22207361 Aleksandra_Olsza
41: 26748 Switzerland
42: 147724 French_Open
43: 91195 The_Championships,_Wimbledon
44: 97851 Australian_Open
45: 19047 Martina_Hingis
46: 534624 Arantxa_Sánchez_Vicario
47: 19047 Martina_Hingis
48: 5462144 Arantxa_Parra_Santonja
49: 22207361 Aleksandra_Olsza
50: 1111175 Magdalena_Maleeva
51: 22207361 Aleksandra_Olsza

52: 22207361 Aleksandra_Olsza
53: 15573 Japan
54: 3484709 Brenda_Schultz-McCarthy
55: 21148 Netherlands
56: 15573 Japan
57: 5321 Czech_Republic
58: 1920271 Helena_Suková
59: 26964606 Austria
60: 3128583 Barbara_Paulus
61: 5405 China

1: 2211473 USTA_Billie_Jean_National_Tennis_Center
[winner: 2211473
USTA_Billie_Jean_National_Tennis_Center, pem_pos=4,
final_scores=0.029256716, final_local_scores=0.088529944,
pem=-1.609438, similarity=15.483996]
gt: 2211473
USTA_Billie_Jean_National_Tennis_Center, pem_pos=4,
final_scores=0.029256716, final_local_scores=0.088529944,
pem=-1.609438, similarity=15.483996]

1: 1443478 Anke_Huber
[pem_pos=0, final_scores=0.8005946,
final_local_scores=0.82366717, pem=0.0, similarity=28.669584]

2: 1111175 Magdalena_Maleeva
[pem_pos=0, final_scores=0.39131325,
final_local_scores=0.3641848, pem=0.0, similarity=14.957713]

3: 645042 New_York_City
[pem_pos=1, final_scores=0.34220856,
final_local_scores=0.4386012, pem=-1.7906687,
similarity=26.916382]

4: 19047 Martina_Hingis
[pem_pos=0, final_scores=0.8405946,
final_local_scores=0.84253633, pem=0.0, similarity=29.232677]

5: 97851 Australian_Open
[pem_pos=0, final_scores=0.38889283,
final_local_scores=0.3604366, pem=-0.3261775,
similarity=16.619663]

6: 1443478 Anke_Huber
[pem_pos=0, final_scores=0.9462963,
final_local_scores=0.9798283, pem=0.0, similarity=33.329742]

7: 1111175 Magdalena_Maleeva
[pem_pos=0, final_scores=0.7353726,
final_local_scores=0.73792946, pem=0.0, similarity=26.111]

8: 82729 US_Open_(tennis)
[pem_pos=1, final_scores=0.5473503,
final_local_scores=0.5354689, pem=-1.2370842,
similarity=26.796627]

9: 19047 Martina_Hingis
[pem_pos=0, final_scores=0.557694,
final_local_scores=0.5335104, pem=0.0, similarity=20.010727]

10: 197638 Grand_Slam_(tennis)
[pem_pos=0, final_scores=0.5593379,
final_local_scores=0.5398726, pem=-0.37449613,

```
similarity=22.237156]
11: 12450241 Ángeles_Montolio
    [pem_pos=0, final_scores=0.4716552,
final_local_scores=0.6072693, pem=0.0, similarity=22.21184]
12: 26667 Spain
    [pem_pos=0, final_scores=0.53562826,
final_local_scores=0.63862526, pem=-0.18857051,
similarity=24.17304]
13: 19047 Martina_Hingis
    [pem_pos=0, final_scores=0.70442814,
final_local_scores=0.6937951, pem=0.0, similarity=24.793941]
14: 963152 Spanish_people
    [pem_pos=0, final_scores=0.22820054,
final_local_scores=0.31325114, pem=-0.22585543,
similarity=14.665988]
15: 1443478 Anke_Huber
    [pem_pos=0, final_scores=0.83794457,
final_local_scores=0.8638994, pem=0.0, similarity=29.870193]
16: 11867 Germany
    [pem_pos=0, final_scores=0.65497035,
final_local_scores=0.75490457, pem=-0.15529114,
similarity=27.462069]
17: 1111175 Magdalena_Maleeva
    [pem_pos=0, final_scores=0.43065327,
final_local_scores=0.40715784, pem=0.0, similarity=16.240114]
18: 3415 Bulgaria
    [pem_pos=0, final_scores=0.66330665,
final_local_scores=0.7948942, pem=-0.31727514,
similarity=29.536333]
19: 1443478 Anke_Huber
    [pem_pos=0, final_scores=0.8580207,
final_local_scores=0.8858296, pem=0.0, similarity=30.524637]
20: 300468 Monica_Seles
    [pem_pos=0, final_scores=0.8521144,
final_local_scores=0.8597789, pem=0.0, similarity=29.74723]
21: 97851 Australian_Open
    [pem_pos=0, final_scores=0.37546813,
final_local_scores=0.34577218, pem=-0.3261775,
similarity=16.182049]
22: 1951841 Amanda_Coetzer
    [pem_pos=0, final_scores=0.785389,
final_local_scores=0.81118995, pem=0.0, similarity=28.29724]
23: 17416221 South_Africa
    [pem_pos=0, final_scores=0.5601778,
final_local_scores=0.7000545, pem=-0.26251382,
similarity=26.408327]
24: 1951841 Amanda_Coetzer
    [pem_pos=0, final_scores=0.4131499,
final_local_scores=0.40457517, pem=0.0, similarity=16.163042]
25: 1443478 Anke_Huber
    [pem_pos=0, final_scores=0.87939376,
final_local_scores=0.9091764, pem=0.0, similarity=31.221352]
```

```
26: 97851 Australian_Open
    [pem_pos=0, final_scores=0.41815507,
final_local_scores=0.39240116, pem=-0.3261775,
similarity=17.57355]
27: 1951841 Amanda_Coetzer
    [pem_pos=0, final_scores=0.5725858,
final_local_scores=0.57873476, pem=0.0, similarity=21.360312]
28: 1443478 Anke_Huber
    [pem_pos=0, final_scores=0.9266395,
final_local_scores=0.9607852, pem=0.0, similarity=32.76146]
29: 107656 Manhattan_Beach,_California
    [pem_pos=0, final_scores=0.25996405,
final_local_scores=0.3691826, pem=-0.37280393,
similarity=17.134224]
30: 152735 Germans
    [pem_pos=2, final_scores=0.15422074,
final_local_scores=0.22734925, pem=-2.1953664,
similarity=22.813011]
31: 197638 Grand_Slam_(tennis)
    [pem_pos=0, final_scores=0.5814198,
final_local_scores=0.56399375, pem=-0.37449613,
similarity=22.95698]
32: 197638 Grand_Slam_(tennis)
    [pem_pos=0, final_scores=0.574038,
final_local_scores=0.55593026, pem=-0.37449613,
similarity=22.716349]
33: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.45405316,
final_local_scores=0.43355596, pem=-1.2370842,
similarity=23.755342]
34: 22207361 Aleksandra_Olsza
    [pem_pos=0, final_scores=0.8574526,
final_local_scores=0.9477754, pem=0.0, similarity=32.373222]
35: 22936 Poland
    [pem_pos=0, final_scores=0.6047415,
final_local_scores=0.69571555, pem=-0.15333635,
similarity=25.68512]
36: 22207361 Aleksandra_Olsza
    [pem_pos=0, final_scores=1.0274733,
final_local_scores=1.1376157, pem=0.0, similarity=38.038437]
37: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.5584171,
final_local_scores=0.5419948, pem=-0.8505555,
similarity=24.889372]
38: 1111175 Magdalena_Maleeva
    [pem_pos=0, final_scores=0.4620059,
final_local_scores=0.44140577, pem=0.0, similarity=17.262142]
39: 19047 Martina_Hingis
    [pem_pos=0, final_scores=0.66005605,
final_local_scores=0.6453253, pem=0.0, similarity=23.347507]
40: 22207361 Aleksandra_Olsza
    [pem_pos=0, final_scores=0.98683125,
```

```
final_local_scores=1.0932202, pem=0.0, similarity=36.71359]
41: 26748 Switzerland
    [pem_pos=0, final_scores=0.6616088,
final_local_scores=0.7593547, pem=-0.1392413,
similarity=27.507587]
42: 147724 French_Open
    [pem_pos=0, final_scores=0.5529104,
final_local_scores=0.54148203, pem=-0.23693618,
similarity=21.537111]
43: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.5054756,
final_local_scores=0.48416424, pem=-0.8505555,
similarity=23.163593]
44: 97851 Australian_Open
    [pem_pos=0, final_scores=0.43614727,
final_local_scores=0.4120549, pem=-0.3261775,
similarity=18.160057]
45: 19047 Martina_Hingis
    [pem_pos=0, final_scores=0.7034401,
final_local_scores=0.6927158, pem=0.0, similarity=24.761734]
46: 534624 Arantxa_Sánchez_Vicario
    [pem_pos=0, final_scores=0.5562188,
final_local_scores=0.548792, pem=0.0, similarity=20.46676]
47: 19047 Martina_Hingis
    [pem_pos=0, final_scores=0.643193,
final_local_scores=0.626905, pem=0.0, similarity=22.79781]
48: 5462144 Arantxa_Parra_Santonja
    [pem_pos=3, final_scores=0.1308137,
final_local_scores=0.12572241, pem=-1.7917595,
similarity=17.585386]
49: 22207361 Aleksandra_Olsza
    [pem_pos=0, final_scores=1.0575705,
final_local_scores=1.1704922, pem=0.0, similarity=39.019535]
50: 1111175 Magdalena_Maleeva
    [pem_pos=0, final_scores=0.5362587,
final_local_scores=0.5225157, pem=0.0, similarity=19.682625]
51: 22207361 Aleksandra_Olsza
    [pem_pos=0, final_scores=1.0889411,
final_local_scores=1.2047598, pem=0.0, similarity=40.04215]
52: 22207361 Aleksandra_Olsza
    [pem_pos=0, final_scores=1.0014305,
final_local_scores=1.1091678, pem=0.0, similarity=37.189495]
53: 15573 Japan
    [pem_pos=0, final_scores=0.69692916,
final_local_scores=0.83416873, pem=-0.99076414,
similarity=34.3709]
54: 3484709 Brenda_Schultz-McCarthy
    [pem_pos=0, final_scores=0.79983014,
final_local_scores=0.86098754, pem=0.0, similarity=29.783298]
55: 21148 Netherlands
    [pem_pos=0, final_scores=0.43413442,
final_local_scores=0.5125402, pem=-0.24940795,
```

```
similarity=20.741255]
56: 15573 Japan
    [pem_pos=0, final_scores=0.68306786,
final_local_scores=0.81902736, pem=-0.13549899,
similarity=29.267986]
57: 5321 Czech_Republic
    [pem_pos=1, final_scores=0.48737246,
final_local_scores=0.5648075, pem=-1.208016,
similarity=27.514072]
58: 1920271 Helena_Suková
    [pem_pos=0, final_scores=0.8674241,
final_local_scores=0.9095827, pem=0.0, similarity=31.233475]
59: 26964606 Austria
    [pem_pos=0, final_scores=0.60771227,
final_local_scores=0.6967909, pem=-0.40674555,
similarity=27.095287]
60: 3128583 Barbara_Paulus
    [pem_pos=0, final_scores=0.6158652,
final_local_scores=0.68651617, pem=0.0, similarity=24.576725]
61: 5405 China
    [pem_pos=0, final_scores=0.39296407,
final_local_scores=0.5122771, pem=-0.15669827,
similarity=20.229235]
```

Sample: b'436_TENNIS'

TENNIS - [1 1 STICH 1] GLAD HE STAYED AFTER OPEN VICTORY .
Richard Finn [2 2 NEW YORK 2] 1996-08-26 [3 3 Michael Stich 3]
nearly pulled out of the [4 4 U.S. Open 4] in protest over
the men 's seeding fiasco but the former [5 5 Wimbledon 5]
champion was glad he stayed after sweating out a four-set win
on Monday over qualifier [6 6 Tommy Haas 6] . " I still feel
it 's embarrassing what happened and I was about to pull out
yesterday and say , ' That 's it , ' " said [7 7 Stich 7] ,
one of a host of men who cried foul over seeding procedures
that forced an unprecedented remaking of the men 's draw last
week . " But there are so many reasons to play , especially
spectators and the kids who come out here and want to enjoy
watching tennis , that I decided to stay . " In a break from
tradition , the Open did not seed in strict accordance with [8
8 ATP 8] rankings , instead taking into account other factors
that raised objections of favouritism toward [9 9 U.S. 9]
players . One prominent player that did not stay for the Open
was [10 10 French Open 10] champion Yvgeny Kafelnikov , who
after being dropped three spots from his [11 11 ATP 11]
ranking to a seventh seeding , withdrew and returned home to [
12 12 Russia 12] . [13 13 Kafelnikov 13] had pulled out of
last week 's tournament with a rib injury . At a news
conference attended by approximately 50 players on Sunday , [1
4 14 U.S. 14] [15 15 Davis Cup 15] player [16 16 Todd Martin 1
6] expressed the players ' outrage at the seedings . " The
way the [17 17 U.S. Open 17] has seeded here , tampering with
the ranking system , has tarnished the image and reputation

of this [18 18 U.S. Open 18] in the players ' mind , and we think that is damaging to our sport , " [19 19 Martin 19] said . [20 20 Stich 20] said he felt the players ought to have organised an active protest . " I feel that we made it a little easy for the [21 21 USTA 21] . They did n't really get hurt as much as I think they should have , " said [22 22 Stich 22] . " I feel that we should have maybe just cancelled out the Monday , not show up today and start the tournament tomorrow . " But once the 27-year-old [23 23 Stich 23] got on the court , he focused his energies on trying to win the year 's last [24 24 Grand Slam 24] . He took a positive first step with his 6-3 1-6 6-1 7-5 win over compatriot [25 25 Haas 25] on a sun-baked Grandstand court . Others advancing early on Monday included 11th-seeded [26 26 American 26] [27 27 Malivai Washington 27] , the [28 28 Wimbledon 28] runner-up , [29 29 Sweden 29] 's [30 30 Magnus Gustafsson 30] , and two-time former [31 31 French Open 31] champion [32 32 Sergi Bruguera 32] of [33 33 Spain 33] , who will be [34 34 Stich 34] 's next opponent . The suspicion , however , lingers in [35 35 Stich 35] 's mind that [36 36 U.S. Open 36] officials did tamper with the seeding process in order to benefit homegrown players . " I get the feeling that everything is done here for the [37 37 American 37] players and they forget about all the other players , " said [38 38 Stich 38] , who lost the 1994 Open final to [39 39 Andre Agassi 39] . It was [40 40 Agassi 40] who was at the centre of the controversy that engulfed the tournament since the original draw was completed on Wednesday . The flamboyant [41 41 American 41] star was bumped up two spots from his [42 42 ATP 42] ranking of eight to a seeding of six . " He ([43 43 Agassi 43]) should be seeded the way he is playing tennis right now , " said [44 44 Stich 44] about the unfairness of moving up [45 45 Agassi 45] , who made early exits from the [46 46 French Open 46] and [47 47 Wimbledon 47] . [48 48 Stich 48] , not seeded here for the first time since 1990 , might have benefitted from some fiddling with the seedings himself after [49 49 Kafelnikov 49] withdrew . Ranked 18th in the world , [50 50 Stich 50] might have been slipped into that spot ahead of [51 51 Spain 51] 's [52 52 Felix Mantilla 52] , who is 16th but had never been played in the Open and had been left out of the seedings originally . But [53 53 Stich 53] did n't want to play that game . " I think he deserves to be seeded as everybody else who is in the top 16 deserves to be seeded , " [54 54 Stich 54] said .

- 1: 979088 Michael_Stich
- 2: 645042 New_York_City
- 3: 979088 Michael_Stich
- 4: 82729 US_Open_(tennis)
- 5: 91195 The_Championships,_Wimbledon
- 6: 1118005 Tommy_Haas
- 7: 979088 Michael_Stich
- 8: 627107 Association_of_Tennis_Professionals

```
9: 3434750 United_States
10: 147724 French_Open
11: 627107 Association_of_Tennis_Professionals
12: 25391 Russia
13: 489963 Yevgeny_Kafelnikov
14: 3434750 United_States
15: 227983 Davis_Cup
16: 871725 Todd_Martin
17: 82729 US_Open_(tennis)
18: 82729 US_Open_(tennis)
19: 871725 Todd_Martin
20: 979088 Michael_Stich
21: 1178132 United_States_Tennis_Association
22: 979088 Michael_Stich
23: 979088 Michael_Stich
24: 197638 Grand_Slam_(tennis)
25: 1118005 Tommy_Haas
26: 3434750 United_States
27: 873616 MaliVai_Washington
28: 91195 The_Championships,_Wimbledon
29: 5058739 Sweden
30: 2793059 Magnus_Gustafsson
31: 147724 French_Open
32: 913625 Sergi_Bruguera
33: 26667 Spain
34: 979088 Michael_Stich
35: 979088 Michael_Stich
36: 82729 US_Open_(tennis)
37: 3434750 United_States
38: 979088 Michael_Stich
39: 595 Andre_Agassi
40: 595 Andre_Agassi
41: 3434750 United_States
42: 627107 Association_of_Tennis_Professionals
43: 595 Andre_Agassi
44: 979088 Michael_Stich
45: 595 Andre_Agassi
46: 147724 French_Open
47: 91195 The_Championships,_Wimbledon
48: 979088 Michael_Stich
49: 489963 Yevgeny_Kafelnikov
50: 979088 Michael_Stich
51: 26667 Spain
52: 1723559 Félix_Mantilla_Botella
53: 979088 Michael_Stich
54: 979088 Michael_Stich
1: 979088 Michael_Stich
    [pem_pos=0, final_scores=0.78198224,
final_local_scores=0.76187915, pem=0.0, similarity=26.825706]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.36172324,
final_local_scores=0.4679821, pem=-1.7906687,
```

```
similarity=27.793167]
3: 979088 Michael_Stich
    [pem_pos=0, final_scores=0.8020206,
final_local_scores=0.7825339, pem=0.0, similarity=27.442085]
4: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.5011349,
final_local_scores=0.4827754, pem=-1.2370842,
similarity=25.22415]
5: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.5546673,
final_local_scores=0.53651816, pem=-0.8505555,
similarity=24.725939]
6: 1118005 Tommy_Haas
    [pem_pos=0, final_scores=0.657903,
final_local_scores=0.64680445, pem=0.0, similarity=23.391647]
7: 979088 Michael_Stich
    [pem_pos=0, final_scores=1.0348997,
final_local_scores=1.0381532, pem=0.0, similarity=35.070274]
8: 627107 Association_of_Tennis_Professionals
    [pem_pos=1, final_scores=0.5049252,
final_local_scores=0.49309176, pem=-0.88842654,
similarity=23.635958]
9: 3434750 United_States
    [pem_pos=0, final_scores=0.40526068,
final_local_scores=0.51218754, pem=-0.14075686,
similarity=20.139868]
10: 147724 French_Open
    [pem_pos=0, final_scores=0.4972586,
final_local_scores=0.47257185, pem=-0.23693618,
similarity=19.480694]
11: 627107 Association_of_Tennis_Professionals
    [pem_pos=1, final_scores=0.4689055,
final_local_scores=0.45374566, pem=-0.88842654,
similarity=22.46179]
12: 25391 Russia
    [pem_pos=0, final_scores=0.56537014,
final_local_scores=0.71269166, pem=-0.20299491,
similarity=26.461771]
13: 489963 Yevgeny_Kafelnikov
    [pem_pos=0, final_scores=0.5571564,
final_local_scores=0.54736775, pem=0.0, similarity=20.42426]
14: 3434750 United_States
    [pem_pos=0, final_scores=0.4014269,
final_local_scores=0.5079997, pem=-0.14075686,
similarity=20.014894]
15: 227983 Davis_Cup
    [pem_pos=0, final_scores=0.34661847,
final_local_scores=0.34004405, pem=-0.21573581,
similarity=15.410511]
16: 871725 Todd_Martin
    [pem_pos=0, final_scores=0.89096934,
final_local_scores=0.91021734, pem=0.0, similarity=31.252415]
```

```
17: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.4928128,
final_local_scores=0.4736848, pem=-1.2370842,
similarity=24.952868]
18: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.4988495,
final_local_scores=0.48027897, pem=-1.2370842,
similarity=25.14965]
19: 871725 Todd_Martin
    [pem_pos=0, final_scores=0.9864488,
final_local_scores=1.0145142, pem=0.0, similarity=34.364838]
20: 979088 Michael_Stich
    [pem_pos=0, final_scores=1.0497934,
final_local_scores=1.0544221, pem=0.0, similarity=35.555775]
21: 1178132 United_States_Tennis_Association
    [pem_pos=0, final_scores=0.3322534,
final_local_scores=0.3561606, pem=0.0, similarity=14.718256]
22: 979088 Michael_Stich
    [pem_pos=0, final_scores=1.1524925,
final_local_scores=1.1666055, pem=0.0, similarity=38.90355]
23: 979088 Michael_Stich
    [pem_pos=0, final_scores=0.97259194,
final_local_scores=0.97009116, pem=0.0, similarity=33.03917]
24: 197638 Grand_Slam_(tennis)
    [pem_pos=0, final_scores=0.57435644,
final_local_scores=0.56066984, pem=-0.37449613,
similarity=22.857788]
25: 1118005 Tommy_Haas
    [pem_pos=0, final_scores=1.0559865,
final_local_scores=1.0835979, pem=0.0, similarity=36.426434]
26: 3434750 United_States
    [pem_pos=0, final_scores=0.52347875,
final_local_scores=0.6413228, pem=-0.12994486,
similarity=23.934725]
27: 873616 MaliVai_Washington
    [pem_pos=0, final_scores=0.7124199,
final_local_scores=0.7522538, pem=0.0, similarity=26.538465]
28: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.5120213,
final_local_scores=0.48993385, pem=-0.8505555,
similarity=23.33577]
29: 5058739 Sweden
    [pem_pos=0, final_scores=0.68614,
final_local_scores=0.8117032, pem=-0.2135508,
similarity=29.473877]
30: 2793059 Magnus_Gustafsson
    [pem_pos=0, final_scores=0.8313076,
final_local_scores=0.8806413, pem=0.0, similarity=30.369806]
31: 147724 French_Open
    [pem_pos=0, final_scores=0.51074487,
final_local_scores=0.4873036, pem=-0.23693618,
similarity=19.92032]
```

32: 913625 Sergi_Bruguera
[pem_pos=0, final_scores=0.81777984,
final_local_scores=0.8365421, pem=0.0, similarity=29.053799]
33: 26667 Spain
[pem_pos=2, final_scores=0.6023414,
final_local_scores=0.729459, pem=-0.18857051,
similarity=26.8837]
34: 979088 Michael_Stich
[pem_pos=0, final_scores=0.80405694,
final_local_scores=0.78599244, pem=0.0, similarity=27.545296]
35: 979088 Michael_Stich
[pem_pos=0, final_scores=0.7677992,
final_local_scores=0.7463863, pem=0.0, similarity=26.363369]
36: 82729 US_Open_(tennis)
[pem_pos=1, final_scores=0.5081108,
final_local_scores=0.49039555, pem=-1.2370842,
similarity=25.45155]
37: 3434750 United_States
[pem_pos=0, final_scores=0.43530375,
final_local_scores=0.545005, pem=-0.12994486,
similarity=21.060411]
38: 979088 Michael_Stich
[pem_pos=0, final_scores=1.0796931,
final_local_scores=1.0870831, pem=0.0, similarity=36.53044]
39: 595 Andre_Agassi
[pem_pos=0, final_scores=0.6702089,
final_local_scores=0.65305996, pem=0.0, similarity=23.578325]
40: 595 Andre_Agassi
[pem_pos=0, final_scores=0.7127698,
final_local_scores=0.69955134, pem=0.0, similarity=24.965721]
41: 3434750 United_States
[pem_pos=0, final_scores=0.30284303,
final_local_scores=0.40031183, pem=-0.12994486,
similarity=16.742477]
42: 627107 Association_of_Tennis_Professionals
[pem_pos=1, final_scores=0.53470635,
final_local_scores=0.52562314, pem=-0.88842654,
similarity=24.606756]
43: 595 Andre_Agassi
[pem_pos=0, final_scores=0.5560891,
final_local_scores=0.52840143, pem=0.0, similarity=19.858265]
44: 979088 Michael_Stich
[pem_pos=0, final_scores=0.993751,
final_local_scores=0.9932043, pem=0.0, similarity=33.72891]
45: 595 Andre_Agassi
[pem_pos=0, final_scores=0.5337835,
final_local_scores=0.504036, pem=0.0, similarity=19.131153]
46: 147724 French_Open
[pem_pos=0, final_scores=0.520592,
final_local_scores=0.49806005, pem=-0.23693618,
similarity=20.241314]
47: 91195 The_Championships,_Wimbledon

```
[pem_pos=0, final_scores=0.5009702,
final_local_scores=0.47786218, pem=-0.8505555,
similarity=22.975527]
48: 979088 Michael_Stich
    [pem_pos=0, final_scores=0.7802122,
final_local_scores=0.7599457, pem=0.0, similarity=26.768007]
49: 489963 Yevgeny_Kafelnikov
    [pem_pos=0, final_scores=0.6197303,
final_local_scores=0.6157203, pem=0.0, similarity=22.464033]
50: 979088 Michael_Stich
    [pem_pos=0, final_scores=1.0154228,
final_local_scores=1.0168777, pem=0.0, similarity=34.435368]
51: 26667 Spain
    [pem_pos=2, final_scores=0.646386,
final_local_scores=0.777571, pem=-0.18857051,
similarity=28.319458]
52: 1723559 Félix_Mantilla_Botella
    [pem_pos=0, final_scores=0.71216315,
final_local_scores=0.7522736, pem=0.0, similarity=26.539059]
53: 979088 Michael_Stich
    [pem_pos=0, final_scores=1.1181974,
final_local_scores=1.1291432, pem=0.0, similarity=37.7856]
54: 979088 Michael_Stich
    [pem_pos=0, final_scores=0.88885385,
final_local_scores=0.8786202, pem=0.0, similarity=30.309494]
```

Sample: b'437_TENNIS'

TENNIS - MONDAY 'S RESULTS FROM [1 1 U.S. OPEN 1] . [2 2 NEW YORK 2] 1996-08-26 Results of first round matches on Monday in the [3 3 U.S. Open 3] tennis championships at the [1 National Tennis Centre 1] (prefix denotes seeding) : Women 's singles 16 - [4 4 Martina Hingis 4] ([5 5 Switzerland 5]) beat [6 6 Angeles Montolio 6] ([7 7 Spain 7]) 6-1 6-0 [8 8 Anne-Gaelle Sidot 8] ([9 9 France 9]) beat [10 10 Janet te Husarova 10] ([11 11 Slovakia 11]) 6-4 6-4 13 - [12 12 Brenda Schultz-McCarthy 12] ([13 13 Netherlands 13]) beat Nana Miyagi ([14 14 Japan 14]) 6-1 6-4 [15 15 Aleksandra Ol sza 15] ([16 16 Poland 16]) beat 12 - [17 17 Magdalena Male eva 17] ([18 18 Bulgaria 18]) 6-4 6-2 Men 's singles [19 19 Michael Stich 19] ([20 20 Germany 20]) beat [21 21 Tommy Ha as 21] ([22 22 Germany 22]) 6-3 1-6 6-1 7-5 [23 23 Sergi Br uguera 23] ([24 24 Spain 24]) beat Kris Goossens ([25 25 Belgium 25]) 6-2 6-0 7-6 (7-1) Frederic Vitoux ([26 26 France 26]) beat [27 27 Ramon Delgado 27] ([28 28 Paraguay 28]) 6-4 6-4 7-6 (7-3) Women 's singles Henrietta Nagyova ([29 29 Slovakia 29]) beat [30 30 Gala Leon Garcia 30] ([31 31 Spain 31]) 6-1 4-6 6-3 [32 32 Asa Carlsson 32] ([33 33 Sweden 33]) beat [34 34 Gloria Pizzichini 34] ([35 35 Italy 35]) 3-6 6-1 7-5 [36 36 Barbara Schett 36] ([37 37 Austria 37]) beat [38 38 Sabine Appelmans 38] ([39 39 Belgium 39]) 1-6 6-4 6-4 Cristina Torrens-Valero ([40 40 Spain 40]) beat [41 41 Sabine Hack 41] ([42 42 Germany 42]) 2-6 6-4 6-2 Women

's singles [43 43 Helena Sukova 43] ([44 44 Czech Republic 44]) beat Yone Kamio ([45 45 Japan 45]) 6-2 6-3 [46 46 Irina Spirlea 46] ([47 47 Romania 47]) beat Petra Begerow ([48 48 Germany 48]) 6-3 6-2 Maria Jose Gaidano ([49 49 Argentina 49]) beat Melanie Schnell ([50 50 Austria 50]) 6-4 6-0 Men 's singles [51 51 Carlos Moya 51] ([52 52 Spain 52]) beat [53 53 Scott Humphries 53] ([54 54 U.S. 54]) 6-1 6-7 (3-7) 6-7 (1-7) 6-0 6-4 [55 55 Kenneth Carlsen 55] ([56 56 Denmark 56]) beat [57 57 Patrick Rafter 57] ([58 58 Australia 58]) 7-6 (9-7) 6-3 7-6 (8-6) [59 59 Magnus Gustafsson 59] ([60 60 Sweden 60]) beat [61 61 Carlos Costa 61] ([62 62 Spain 62]) 7-5 4-6 7-6 (7-4) 6-3 [63 63 Jeff Tarango 63] ([64 64 U.S. 64]) beat [65 65 Alex Radulescu 65] ([66 66 Romania 66]) 6-7 (5-7) 6-4 6-1 retired , heat exhaustion Men 's singles 11 - [67 67 MaliVai Washington 67] ([68 68 U.S. 68]) beat [69 69 Karim Alami 69] ([70 70 Morocco 70]) 6-4 2-6 7-6 (7-5) 6-1 Dirk Dier ([71 71 Germany 71]) beat [72 72 Chuck Adams 72] ([73 73 U.S. 73]) 6-4 2-6 6-4 6-4 [74 74 Jason Stoltenberg 74] ([75 75 Australia 75]) beat [76 76 Stefano Pescosolido 76] ([77 77 Italy 77]) 7-5 6-4 6-1 [78 78 Arnaud Boetsch 78] ([79 79 France 79]) beat [80 80 Nicolas Pereira 80] ([81 81 Venezuela 81]) 7-6 (7-4) 6-4 7-5 [82 82 David Prinosil 82] ([83 83 Germany 83]) beat [84 84 Peter Tramacchi 84] ([85 85 Australia 85]) 6-3 6-2 3-6 6-7 (5-7) 6-1 Women 's singles [86 86 Amanda Coetzer 86] ([87 87 South Africa 87]) beat 6 - [88 88 Anke Huber 88] ([89 89 Germany 89]) 6-1 2-6 6-2 [90 90 Anna Kournikova 90] ([91 91 Russia 91]) beat [92 92 Ludmila Richterova 92] ([93 93 Czech Republic 93]) 7-6 (7-4) 6-3 [94 94 Debbie Graham 94] ([95 95 U.S. 95]) beat Stephanie Deville ([96 96 Belarus 96]) 6-4 6-2 [97 97 Barbara Rittner 97] ([98 98 Germany 98]) beat [99 99 Katarina Studenikova 99] ([100 100 Slovakia 100]) 7-5 7-5 [101 101 Kristina Brandi 101] ([102 102 U.S. 102]) beat [103 103 Andrea Glass 103] ([104 104 Germany 104]) 6-2 6-3 [105 105 Ines Gorrochategui 105] ([106 106 Argentina 106]) beat [107 107 Magdalena Grzybowska 107] ([108 108 Poland 108]) 4-6 6-4 6-1 Men 's singles [109 109 Alberto Berasategui 109] ([110 110 Spain 110]) beat [111 111 Cecil Mamiit 111] ([112 112 U.S. 112]) 6-1 6-4 6-0 [113 113 Guillaume Raoux 113] ([114 114 France 114]) beat [115 115 Filip Dewulf 115] ([116 116 Belgium 116]) 7-6 (7-5) 3-6 1-6 6-4 7-5 Alex O'Brien ([117 117 U.S. 117]) beat [118 118 Nicolas Lapentti 118] ([119 119 Ecuador 119]) 6-4 1-6 6-4 6-3 2 - [120 120 Michael Chang 120] ([121 121 U.S. 121]) beat [122 122 Jaime Oncins 122] ([123 123 Brazil 123]) 3-6 6-1 6-0 7 - 6 (8-6) Women 's singles 14 - [124 124 Barbara Paulus 124] ([125 125 Austria 125]) beat Yi Jing-Qian ([126 126 China 126]) 6-2 6-1 Wang Shi-Ting ([127 127 Taiwan 127]) beat [128 128 Corina Morariu 128] ([129 129 U.S. 129]) 6-4 6-7 (5-7) 6-2 [130 130 Linda Wild 130] ([131 131 U.S. 131]) beat [132 132 Sung-Hee Park 132] ([133 133 South Korea 133]) 6-2 6-3 [134 134 Sarah Pitkowski 134] ([135 135 France 135]) beat [136 136 Meghann Shaughnessy 136

] ([137 137 U.S. 137]) 6-3 6- 3 [138 138 Dally Randriantefy
138] ([139 139 Madagascar 139]) beat [140 140 Elena Makarova
140] ([141 141 Russia 141]) 6- 3 1-6 7-5 [142 142 Laurence C
ourtois 142] ([143 143 Belgium 143]) beat Flora Perfetti ([
144 144 Italy 144]) 6-4 3-6 6-2 Men 's singles [145 145 Lean
der Paes 145] ([146 146 India 146]) beat [147 147 Marcos On
druska 147] ([148 148 South Africa 148]) 7-6 (7-3) 6-2
7-5 [149 149 Jan Siemerink 149] ([150 150 Netherlands 150])
beat [151 151 Carl-Uwe Steeb 151] ([152 152 Germany 152])
4-6 6 - 1 7-6 (7-4) 6-4 [153 153 Neville Godwin 153] ([154
154 South Africa 154]) beat [155 155 Tomas Carbonell 155] ([
156 156 Spain 156]) 6-4 6-2 3-6 6-1 [157 157 Jim Grabb 157]
([158 158 U.S. 158]) beat [159 159 Sandon Stolle 159] ([160
160 Australia 160]) 6-3 7-5 7-6 (7-4) Women 's singles [161
161 Alexandra Fusai 161] ([162 162 France 162]) beat [163 16
3 Jill Craybas 163] ([164 164 U.S. 164]) 6-1 2-6 7-5 Naoko
Kijimuta ([165 165 Japan 165]) beat Tatyana Jecmenica ([166
166 Yugoslavia 166]) 6-3 6-2 [167 167 Nathalie Dechy 167] ([
168 168 France 168]) beat Christina Singer ([169 169 Germany
169]) 6-4 6-0 Jane Chi ([170 170 U.S. 170]) beat Maria
Antonio Sanchez Lorenzo ([171 171 Spain 171]) 6-4 1-6 6-3 [1
72 172 Els Callens 172] ([173 173 Belgium 173]) beat [174 17
4 Nicole Bradtke 174] ([175 175 Australia 175]) 7-6 (7-1)
7-6 (9-7) Natalia Baudone ([176 176 Italy 176]) beat
Jolene Watanabe ([177 177 U.S. 177]) 6-4 4-6 7-6 (8-6) [17
8 178 Ai Sugiyama 178] ([179 179 Japan 179]) beat [180 180 J
ana Kandarr 180] ([181 181 Germany 181]) 6-2 6-1

- 1: 82729 US_Open_(tennis)
- 2: 645042 New_York_City
- 3: 82729 US_Open_(tennis)
- 4: 19047 Martina_Hingis
- 5: 26748 Switzerland
- 6: 12450241 Angeles_Montolio
- 7: 26667 Spain
- 8: 4450757 Anne-Gaëlle_Sidot
- 9: 5843419 France
- 10: 3845884 Janette_Husárová
- 11: 26830 Slovakia
- 12: 3484709 Brenda_Schultz-McCarthy
- 13: 21148 Netherlands
- 14: 15573 Japan
- 15: 22207361 Aleksandra_Olsza
- 16: 22936 Poland
- 17: 1111175 Magdalena_Maleeva
- 18: 3415 Bulgaria
- 19: 979088 Michael_Stich
- 20: 11867 Germany
- 21: 1118005 Tommy_Haas
- 22: 11867 Germany
- 23: 913625 Sergi_Bruguera
- 24: 26667 Spain
- 25: 3343 Belgium

26: 5843419 France
27: 4027836 Ramón_Delgado
28: 23420 Paraguay
29: 26830 Slovakia
30: 6498487 Gala_León_García
31: 26667 Spain
32: 4038174 Åsa_Svensson
33: 5058739 Sweden
34: 26375250 Gloria_Pizzichini
35: 14532 Italy
36: 2516435 Barbara_Schett
37: 26964606 Austria
38: 1729072 Sabine_Appelmans
39: 3343 Belgium
40: 26667 Spain
41: 15035484 Sabine_Hack
42: 11867 Germany
43: 1920271 Helena_Suková
44: 5321 Czech_Republic
45: 15573 Japan
46: 4038084 Irina_Spîrlea
47: 25445 Romania
48: 11867 Germany
49: 18951905 Argentina
50: 26964606 Austria
51: 669421 Carlos_Moyá
52: 26667 Spain
53: 20867776 Scott_Humphries
54: 3434750 United_States
55: 1804967 Kenneth_Carlsen
56: 76972 Denmark
57: 23297 Patrick_Rafter
58: 4689264 Australia
59: 2793059 Magnus_Gustafsson
60: 5058739 Sweden
61: 2785107 Carlos_Costa_(tennis)
62: 26667 Spain
63: 1862751 Jeff_Tarango
64: 3434750 United_States
65: 4026041 Alex_Rădulescu
66: 25445 Romania
67: 873616 MaliVai_Washington
68: 3434750 United_States
69: 3930942 Karim_Alami
70: 19291 Morocco
71: 11867 Germany
72: 20930737 Chuck_Adams
73: 3434750 United_States
74: 2598440 Jason_Stoltenberg
75: 4689264 Australia
76: 3930788 Stefano_Pescosolido
77: 14532 Italy

78: 3930278 Arnaud_Boetsch
79: 5843419 France
80: 3929855 Nicolás_Pereira
81: 32374 Venezuela
82: 3930698 David_Prinosil
83: 11867 Germany
84: 22881243 Peter_Tramacchi
85: 4689264 Australia
86: 1951841 Amanda_Coetzer
87: 17416221 South_Africa
88: 1443478 Anke_Huber
89: 11867 Germany
90: 890 Anna_Kournikova
91: 25391 Russia
92: 27682452 Ludmila_Richterová
93: 5321 Czech_Republic
94: 12556787 Debbie_Graham
95: 3434750 United_States
96: 3457 Belarus
97: 12668474 Barbara_Rittner
98: 11867 Germany
99: 15290332 Katarína_Studeníková
100: 26830 Slovakia
101: 1301702 Kristina_Brandi
102: 3434750 United_States
103: 15290458 Andrea_Glass
104: 11867 Germany
105: 6501697 Inés_Gorrochategui
106: 18951905 Argentina
107: 4542840 Magdalena_Grzybowska
108: 22936 Poland
109: 1728009 Alberto_Berasategui
110: 26667 Spain
111: 10980025 Cecil_Mamiit
112: 3434750 United_States
113: 3930324 Guillaume_Raoux
114: 5843419 France
115: 3004535 Filip_Dewulf
116: 3343 Belgium
117: 3434750 United_States
118: 2817661 Nicolás_Lapentti
119: 9334 Ecuador
120: 391834 Michael_Chang
121: 3434750 United_States
122: 3931197 Jaime_Oncins
123: 3383 Brazil
124: 3128583 Barbara_Paulus
125: 26964606 Austria
126: 5405 China
127: 25734 Taiwan
128: 2629651 Corina_Morariu
129: 3434750 United_States

130: 15442936 Linda_Wild
131: 3434750 United_States
132: 15339834 Park_Sung-hee
133: 27019 South_Korea
134: 3127732 Sarah_Pitkowski-Malcor
135: 5843419 France
136: 1547567 Meghann_Shaughnessy
137: 3434750 United_States
138: 2620429 Dally_Randriantefy
139: 18964 Madagascar
140: 20877000 Elena_Makarova
141: 25391 Russia
142: 3004436 Laurence_Courtois
143: 3343 Belgium
144: 14532 Italy
145: 37491848 Leander_Paes
146: 14533 India
147: 3930413 Marcos_Ondruska
148: 17416221 South_Africa
149: 2876758 Jan_Siemerink
150: 21148 Netherlands
151: 3254714 Carl-Uwe_Staab
152: 11867 Germany
153: 4026099 Neville_Godwin
154: 17416221 South_Africa
155: 13833030 Tomás_Carbonell
156: 26667 Spain
157: 1915197 Jim_Grabb
158: 3434750 United_States
159: 3823337 Sandon_Stolle
160: 4689264 Australia
161: 4224688 Alexandra_Fusai
162: 5843419 France
163: 2442075 Jill_Craybas
164: 3434750 United_States
165: 15573 Japan
166: 34244 Yugoslavia
167: 1953990 Nathalie_Dechy
168: 5843419 France
169: 11867 Germany
170: 3434750 United_States
171: 26667 Spain
172: 1944281 Els_Callens
173: 3343 Belgium
174: 11642302 Nicole_Bradtke
175: 4689264 Australia
176: 14532 Italy
177: 3434750 United_States
178: 1514559 Ai_Sugiyama
179: 15573 Japan
180: 11510693 Jana_Kandarr
181: 11867 Germany

```
1: 82729 US_Open_(tennis)
   [pem_pos=1, final_scores=0.52184546,
final_local_scores=0.5462779, pem=-1.2370842,
similarity=27.11919]
2: 645042 New_York_City
   [pem_pos=1, final_scores=0.42987347,
final_local_scores=0.4590788, pem=-1.7906687,
similarity=27.527473]
3: 82729 US_Open_(tennis)
   [pem_pos=1, final_scores=0.50560206,
final_local_scores=0.5285344, pem=-1.2370842,
similarity=26.58969]
4: 19047 Martina_Hingis
   [pem_pos=0, final_scores=0.8114286,
final_local_scores=0.8891695, pem=0.0, similarity=30.624304]
5: 26748 Switzerland
   [pem_pos=0, final_scores=0.7480549,
final_local_scores=0.7553706, pem=-0.18598008,
similarity=27.642866]
6: 12450241 Ángeles_Montolio
   [pem_pos=0, final_scores=0.586909,
final_local_scores=0.7392728, pem=0.0, similarity=26.151087]
7: 26667 Spain
   [pem_pos=0, final_scores=0.6510094,
final_local_scores=0.65479755, pem=-0.18857051,
similarity=24.655653]
8: 4450757 Anne-Gaëlle_Sidot
   [pem_pos=0, final_scores=0.53920287,
final_local_scores=0.66326624, pem=0.0, similarity=23.8829]
9: 5843419 France
   [pem_pos=0, final_scores=0.74510145,
final_local_scores=0.74906886, pem=-0.16240834,
similarity=27.326622]
10: 3845884 Janette_Husárová
    [pem_pos=0, final_scores=0.9134434,
final_local_scores=1.0432303, pem=0.0, similarity=35.221786]
11: 26830 Slovakia
    [pem_pos=0, final_scores=0.65484124,
final_local_scores=0.65516526, pem=-0.36602598,
similarity=25.631659]
12: 3484709 Brenda_Schultz-McCarthy
    [pem_pos=0, final_scores=0.6822738,
final_local_scores=0.7986605, pem=0.0, similarity=27.923336]
13: 21148 Netherlands
    [pem_pos=0, final_scores=0.645431,
final_local_scores=0.64675593, pem=-0.24940795,
similarity=24.74652]
14: 15573 Japan
    [pem_pos=0, final_scores=0.74300736,
final_local_scores=0.78701156, pem=-0.13549899,
similarity=28.312572]
15: 22207361 Aleksandra_Olsza
```

[pem_pos=0, final_scores=0.5976493,
final_local_scores=0.7507255, pem=0.0, similarity=26.492859]
16: 22936 Poland
[pem_pos=0, final_scores=0.6105069,
final_local_scores=0.60888237, pem=-0.15333635,
similarity=23.093843]
17: 1111175 Magdalena_Maleeva
[pem_pos=0, final_scores=0.7125842,
final_local_scores=0.79533577, pem=0.0, similarity=27.82412]
18: 3415 Bulgaria
[pem_pos=0, final_scores=0.69998413,
final_local_scores=0.73123604, pem=-0.31727514,
similarity=27.636646]
19: 979088 Michael_Stich
[pem_pos=0, final_scores=0.7834783,
final_local_scores=0.88720137, pem=0.0, similarity=30.56557]
20: 11867 Germany
[pem_pos=0, final_scores=0.8888145,
final_local_scores=0.90734524, pem=-0.15529114,
similarity=32.011204]
21: 1118005 Tommy_Haas
[pem_pos=0, final_scores=0.76851255,
final_local_scores=0.84655493, pem=0.0, similarity=29.3526]
22: 11867 Germany
[pem_pos=0, final_scores=0.8481608,
final_local_scores=0.86293715, pem=-0.15529114,
similarity=30.685976]
23: 913625 Sergi_Bruguera
[pem_pos=0, final_scores=0.7899092,
final_local_scores=0.8806316, pem=0.0, similarity=30.369518]
24: 26667 Spain
[pem_pos=0, final_scores=0.6647087,
final_local_scores=0.6697619, pem=-0.18857051,
similarity=25.102219]
25: 3343 Belgium
[pem_pos=0, final_scores=0.742833,
final_local_scores=0.7379882, pem=-0.24180195,
similarity=27.42771]
26: 5843419 France
[pem_pos=0, final_scores=0.7660461,
final_local_scores=0.7719478, pem=-0.16240834,
similarity=28.009377]
27: 4027836 Ramón_Delgado
[pem_pos=0, final_scores=0.62124497,
final_local_scores=0.72390294, pem=0.0, similarity=25.692421]
28: 23420 Paraguay
[pem_pos=0, final_scores=0.49255383,
final_local_scores=0.4992196, pem=-0.40851754,
similarity=21.209003]
29: 26830 Slovakia
[pem_pos=0, final_scores=0.69582945,
final_local_scores=0.6999387, pem=-0.36602598,

```
similarity=26.967787]
30: 6498487 Gala_León_García
    [pem_pos=0, final_scores=0.45734978,
final_local_scores=0.65974, pem=0.0, similarity=23.77767]
31: 26667 Spain
    [pem_pos=0, final_scores=0.67112225,
final_local_scores=0.6767678, pem=-0.18857051,
similarity=25.31129]
32: 4038174 Åsa_Svensson
    [pem_pos=0, final_scores=0.86459976,
final_local_scores=1.0118563, pem=0.0, similarity=34.285522]
33: 5058739 Sweden
    [pem_pos=0, final_scores=0.7923357,
final_local_scores=0.8159576, pem=-0.2135508,
similarity=29.600838]
34: 26375250 Gloria_Pizzichini
    [pem_pos=0, final_scores=0.6307359,
final_local_scores=0.73890233, pem=0.0, similarity=26.140032]
35: 14532 Italy
    [pem_pos=0, final_scores=0.7584041,
final_local_scores=0.79355043, pem=-0.20612453,
similarity=28.891779]
36: 2516435 Barbara_Schett
    [pem_pos=0, final_scores=0.8596938,
final_local_scores=0.96863776, pem=0.0, similarity=32.995796]
37: 26964606 Austria
    [pem_pos=0, final_scores=0.6554434,
final_local_scores=0.65020543, pem=-0.25262418,
similarity=24.86695]
38: 1729072 Sabine_Appelmans
    [pem_pos=0, final_scores=0.6813064,
final_local_scores=0.78920764, pem=0.0, similarity=27.641243]
39: 3343 Belgium
    [pem_pos=0, final_scores=0.71600723,
final_local_scores=0.70868504, pem=-0.24180195,
similarity=26.553246]
40: 26667 Spain
    [pem_pos=0, final_scores=0.6866779,
final_local_scores=0.6937599, pem=-0.18857051,
similarity=25.818369]
41: 15035484 Sabine_Hack
    [pem_pos=0, final_scores=0.57752407,
final_local_scores=0.66530377, pem=0.0, similarity=23.943705]
42: 11867 Germany
    [pem_pos=0, final_scores=0.8348256,
final_local_scores=0.8483705, pem=-0.15529114,
similarity=30.25128]
43: 1920271 Helena_Suková
    [pem_pos=0, final_scores=0.8027118,
final_local_scores=0.90563816, pem=0.0, similarity=31.11576]
44: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.60685605,
```

```
final_local_scores=0.60418946, pem=-0.27794555,
similarity=23.631443]
45: 15573 Japan
    [pem_pos=0, final_scores=0.7088669,
final_local_scores=0.7497182, pem=-0.13549899,
similarity=27.199663]
46: 4038084 Irina_Spirlea
    [pem_pos=0, final_scores=0.86281174,
final_local_scores=0.9895908, pem=0.0, similarity=33.62108]
47: 25445 Romania
    [pem_pos=0, final_scores=0.74484926,
final_local_scores=0.76094186, pem=-0.21241479,
similarity=27.95288]
48: 11867 Germany
    [pem_pos=0, final_scores=0.8529386,
final_local_scores=0.86815625, pem=-0.15529114,
similarity=30.841724]
49: 18951905 Argentina
    [pem_pos=0, final_scores=0.6703837,
final_local_scores=0.67957306, pem=-0.30003807,
similarity=26.001183]
50: 26964606 Austria
    [pem_pos=0, final_scores=0.6438021,
final_local_scores=0.6374891, pem=-0.25262418,
similarity=24.487469]
51: 669421 Carlos_Moyá
    [pem_pos=0, final_scores=0.7087962,
final_local_scores=0.7845489, pem=0.0, similarity=27.502216]
52: 26667 Spain
    [pem_pos=0, final_scores=0.6345373,
final_local_scores=0.6368042, pem=-0.18857051,
similarity=24.118696]
53: 20867776 Scott_Humphries
    [pem_pos=0, final_scores=0.6553039,
final_local_scores=0.77466196, pem=0.0, similarity=27.20717]
54: 3434750 United_States
    [pem_pos=0, final_scores=0.534988,
final_local_scores=0.5336203, pem=-0.14075686,
similarity=20.779465]
55: 1804967 Kenneth_Carlsen
    [pem_pos=0, final_scores=0.7590223,
final_local_scores=0.86694604, pem=0.0, similarity=29.961111]
56: 76972 Denmark
    [pem_pos=0, final_scores=0.6141762,
final_local_scores=0.6319431, pem=-0.26762357,
similarity=24.403534]
57: 23297 Patrick_Rafter
    [pem_pos=0, final_scores=0.93435127,
final_local_scores=1.0377052, pem=0.0, similarity=35.056908]
58: 4689264 Australia
    [pem_pos=0, final_scores=0.7510099,
final_local_scores=0.7850189, pem=-0.17168477,
```

```
similarity=28.449892]
59: 2793059 Magnus_Gustafsson
    [pem_pos=0, final_scores=0.7986575,
final_local_scores=0.89211696, pem=0.0, similarity=30.712263]
60: 5058739 Sweden
    [pem_pos=0, final_scores=0.7846504,
final_local_scores=0.8075626, pem=-0.2135508,
similarity=29.350313]
61: 2785107 Carlos_Costa_(tennis)
    [pem_pos=0, final_scores=0.34415585,
final_local_scores=0.40048862, pem=-0.021202235,
similarity=16.156393]
62: 26667 Spain
    [pem_pos=0, final_scores=0.6421626,
final_local_scores=0.6451338, pem=-0.18857051,
similarity=24.367268]
63: 1862751 Jeff_Tarango
    [pem_pos=0, final_scores=0.55324215,
final_local_scores=0.6438382, pem=0.0, similarity=23.30313]
64: 3434750 United_States
    [pem_pos=0, final_scores=0.51348025,
final_local_scores=0.5101264, pem=-0.14075686,
similarity=20.078362]
65: 4026041 Alex_Rădulescu
    [pem_pos=0, final_scores=0.6299186,
final_local_scores=0.753952, pem=0.0, similarity=26.589146]
66: 25445 Romania
    [pem_pos=0, final_scores=0.7821683,
final_local_scores=0.8017072, pem=-0.21241479,
similarity=29.1694]
67: 873616 MaliVai_Washington
    [pem_pos=0, final_scores=0.94088084,
final_local_scores=1.037219, pem=0.0, similarity=35.0424]
68: 3434750 United_States
    [pem_pos=0, final_scores=0.52709454,
final_local_scores=0.5249979, pem=-0.14075686,
similarity=20.522154]
69: 3930942 Karim_Alami
    [pem_pos=0, final_scores=0.6589659,
final_local_scores=0.77773374, pem=0.0, similarity=27.29884]
70: 19291 Morocco
    [pem_pos=0, final_scores=0.6729503,
final_local_scores=0.7118805, pem=-0.29598162,
similarity=26.943241]
71: 11867 Germany
    [pem_pos=0, final_scores=0.8785444,
final_local_scores=0.8961267, pem=-0.15529114,
similarity=31.67642]
72: 20930737 Chuck_Adams
    [pem_pos=0, final_scores=0.62284124,
final_local_scores=0.7298508, pem=0.0, similarity=25.869917]
73: 3434750 United_States
```

```
[pem_pos=0, final_scores=0.52885294,
final_local_scores=0.5269187, pem=-0.14075686,
similarity=20.579477]
74: 2598440 Jason_Stoltzenberg
[pem_pos=0, final_scores=0.6503971,
final_local_scores=0.7372567, pem=0.0, similarity=26.090923]
75: 4689264 Australia
[pem_pos=0, final_scores=0.6911419,
final_local_scores=0.7196222, pem=-0.17168477,
similarity=26.498323]
76: 3930788 Stefano_Pescosolido
[pem_pos=0, final_scores=0.6891715,
final_local_scores=0.80322886, pem=0.0, similarity=28.059664]
77: 14532 Italy
[pem_pos=0, final_scores=0.7160246,
final_local_scores=0.7472572, pem=-0.20612453,
similarity=27.510296]
78: 3930278 Arnaud_Boetsch
[pem_pos=0, final_scores=0.66788065,
final_local_scores=0.77767986, pem=0.0, similarity=27.297232]
79: 5843419 France
[pem_pos=0, final_scores=0.7697157,
final_local_scores=0.7759562, pem=-0.16240834,
similarity=28.128996]
80: 3929855 Nicolás_Pereira
[pem_pos=0, final_scores=0.56222785,
final_local_scores=0.6464519, pem=0.0, similarity=23.381126]
81: 32374 Venezuela
[pem_pos=0, final_scores=0.51619464,
final_local_scores=0.52949095, pem=-0.21226841,
similarity=21.045128]
82: 3930698 David_Prinosil
[pem_pos=0, final_scores=0.642715,
final_local_scores=0.7464349, pem=0.0, similarity=26.36482]
83: 11867 Germany
[pem_pos=0, final_scores=0.9226201,
final_local_scores=0.9442728, pem=-0.15529114,
similarity=33.113194]
84: 22881243 Peter_Tramacchi
[pem_pos=0, final_scores=0.5478135,
final_local_scores=0.6756952, pem=0.0, similarity=24.253805]
85: 4689264 Australia
[pem_pos=0, final_scores=0.7455264,
final_local_scores=0.779029, pem=-0.17168477,
similarity=28.271141]
86: 1951841 Amanda_Coetzer
[pem_pos=0, final_scores=0.78894407,
final_local_scores=0.8866343, pem=0.0, similarity=30.548647]
87: 17416221 South_Africa
[pem_pos=0, final_scores=0.68336135,
final_local_scores=0.7341484, pem=-0.26251382,
similarity=27.425756]
```

88: 1443478 Anke_Huber
[pem_pos=0, final_scores=0.84770757,
final_local_scores=0.96764463, pem=0.0, similarity=32.96616]
89: 11867 Germany
[pem_pos=0, final_scores=0.86318713,
final_local_scores=0.8793513, pem=-0.15529114,
similarity=31.175808]
90: 890 Anna_Kournikova
[pem_pos=0, final_scores=0.70725197,
final_local_scores=0.786024, pem=0.0, similarity=27.546236]
91: 25391 Russia
[pem_pos=0, final_scores=0.8638547,
final_local_scores=0.9168251, pem=-0.20299491,
similarity=32.55352]
92: 27682452 Ludmila_Richterová
[pem_pos=0, final_scores=0.48128307,
final_local_scores=0.6113517, pem=0.0, similarity=22.333668]
93: 5321 Czech_Republic
[pem_pos=0, final_scores=0.6516723,
final_local_scores=0.6531445, pem=-0.27794555,
similarity=25.092358]
94: 12556787 Debbie_Graham
[pem_pos=0, final_scores=0.5992891,
final_local_scores=0.7457363, pem=0.0, similarity=26.343971]
95: 3434750 United_States
[pem_pos=0, final_scores=0.5022591,
final_local_scores=0.49786896, pem=-0.14075686,
similarity=19.712572]
96: 3457 Belarus
[pem_pos=0, final_scores=0.62332696,
final_local_scores=0.6383767, pem=-0.4220667,
similarity=25.435411]
97: 12668474 Barbara_Rittner
[pem_pos=0, final_scores=0.6579728,
final_local_scores=0.76484, pem=0.0, similarity=26.914064]
98: 11867 Germany
[pem_pos=0, final_scores=0.92509526,
final_local_scores=0.9469765, pem=-0.15529114,
similarity=33.19388]
99: 15290332 Katarína_Studeníková
[pem_pos=0, final_scores=0.61519593,
final_local_scores=0.7689428, pem=0.0, similarity=27.036499]
100: 26830 Slovakia
[pem_pos=0, final_scores=0.6374132,
final_local_scores=0.63612777, pem=-0.36602598,
similarity=25.063541]
101: 1301702 Kristina_Brandi
[pem_pos=0, final_scores=0.6175148,
final_local_scores=0.7094065, pem=0.0, similarity=25.259817]
102: 3434750 United_States
[pem_pos=0, final_scores=0.5146292,
final_local_scores=0.51138145, pem=-0.14075686,

```
similarity=20.115814]
103: 15290458 Andrea_Glass
    [pem_pos=0, final_scores=0.4846018,
final_local_scores=0.6158063, pem=0.0, similarity=22.4666]
104: 11867 Germany
    [pem_pos=0, final_scores=0.84816223,
final_local_scores=0.8629388, pem=-0.15529114,
similarity=30.686026]
105: 6501697 In  s_Gorrochategui
    [pem_pos=0, final_scores=0.53025454,
final_local_scores=0.66454464, pem=0.0, similarity=23.921051]
106: 18951905 Argentina
    [pem_pos=0, final_scores=0.67083544,
final_local_scores=0.6800665, pem=-0.30003807,
similarity=26.015907]
107: 4542840 Magdalena_Grzybowska
    [pem_pos=0, final_scores=0.52552086,
final_local_scores=0.65124166, pem=0.0, similarity=23.524063]
108: 22936 Poland
    [pem_pos=0, final_scores=0.59987915,
final_local_scores=0.5972732, pem=-0.15333635,
similarity=22.747402]
109: 1728009 Alberto_Berasategui
    [pem_pos=0, final_scores=0.7060037,
final_local_scores=0.7893967, pem=0.0, similarity=27.646885]
110: 26667 Spain
    [pem_pos=0, final_scores=0.6571638,
final_local_scores=0.66152036, pem=-0.18857051,
similarity=24.856276]
111: 10980025 Cecil_Mamiit
    [pem_pos=0, final_scores=0.73725003,
final_local_scores=0.8596879, pem=0.0, similarity=29.744514]
112: 3434750 United_States
    [pem_pos=0, final_scores=0.54370683,
final_local_scores=0.5431443, pem=-0.14075686,
similarity=21.06368]
113: 3930324 Guillaume_Raoux
    [pem_pos=0, final_scores=0.7046905,
final_local_scores=0.79914016, pem=0.0, similarity=27.937649]
114: 5843419 France
    [pem_pos=0, final_scores=0.74358714,
final_local_scores=0.7474147, pem=-0.16240834,
similarity=27.27726]
115: 3004535 Filip_Dewulf
    [pem_pos=0, final_scores=0.68003553,
final_local_scores=0.76380223, pem=0.0, similarity=26.883095]
116: 3343 Belgium
    [pem_pos=0, final_scores=0.71578044,
final_local_scores=0.7084373, pem=-0.24180195,
similarity=26.545853]
117: 3434750 United_States
    [pem_pos=0, final_scores=0.51643926,
```

```
final_local_scores=0.51335865, pem=-0.14075686,
similarity=20.174816]
118: 2817661 Nicolás_Lapentti
    [pem_pos=0, final_scores=0.7421831,
final_local_scores=0.814245, pem=0.0, similarity=28.388407]
119: 9334 Ecuador
    [pem_pos=0, final_scores=0.5171451,
final_local_scores=0.57450765, pem=-0.21404564,
similarity=22.39818]
120: 391834 Michael_Chang
    [pem_pos=0, final_scores=0.7877889,
final_local_scores=0.8546443, pem=0.0, similarity=29.594004]
121: 3434750 United_States
    [pem_pos=0, final_scores=0.50750035,
final_local_scores=0.5035942, pem=-0.14075686,
similarity=19.883427]
122: 3931197 Jaime_Oncins
    [pem_pos=0, final_scores=0.68014604,
final_local_scores=0.7747517, pem=0.0, similarity=27.20985]
123: 3383 Brazil
    [pem_pos=0, final_scores=0.839474,
final_local_scores=0.8741693, pem=-0.21313342,
similarity=31.335724]
124: 3128583 Barbara_Paulus
    [pem_pos=0, final_scores=0.63376033,
final_local_scores=0.76152617, pem=0.0, similarity=26.815172]
125: 26964606 Austria
    [pem_pos=0, final_scores=0.6640002,
final_local_scores=0.65955245, pem=-0.25262418,
similarity=25.145884]
126: 5405 China
    [pem_pos=0, final_scores=0.43275505,
final_local_scores=0.45269758, pem=-0.15669827,
similarity=18.451262]
127: 25734 Taiwan
    [pem_pos=0, final_scores=0.34580016,
final_local_scores=0.37226063, pem=-0.077886656,
similarity=15.6222725]
128: 2629651 Corina_Morariu
    [pem_pos=0, final_scores=0.716355,
final_local_scores=0.83663815, pem=0.0, similarity=29.056664]
129: 3434750 United_States
    [pem_pos=0, final_scores=0.5301885,
final_local_scores=0.5283776, pem=-0.14075686,
similarity=20.623013]
130: 15442936 Linda_Wild
    [pem_pos=0, final_scores=0.66348296,
final_local_scores=0.7915296, pem=0.0, similarity=27.710535]
131: 3434750 United_States
    [pem_pos=0, final_scores=0.52060676,
final_local_scores=0.517911, pem=-0.14075686,
similarity=20.310669]
```

132: 15339834 Park_Sung-hee
[pem_pos=0, final_scores=0.5325815,
final_local_scores=0.6562981, pem=0.0, similarity=23.674957]
133: 27019 South_Korea
[pem_pos=0, final_scores=0.5267819,
final_local_scores=0.5490276, pem=-0.22449642,
similarity=21.69464]
134: 3127732 Sarah_Pitkowski-Malcor
[pem_pos=0, final_scores=0.6383392,
final_local_scores=0.7483128, pem=0.0, similarity=26.420858]
135: 5843419 France
[pem_pos=0, final_scores=0.6990748,
final_local_scores=0.69879174, pem=-0.16240834,
similarity=25.826254]
136: 1547567 Meghann_Shaughnessy
[pem_pos=0, final_scores=0.72615266,
final_local_scores=0.83220077, pem=0.0, similarity=28.924244]
137: 3434750 United_States
[pem_pos=0, final_scores=0.5189951,
final_local_scores=0.51615053, pem=-0.14075686,
similarity=20.258131]
138: 2620429 Dally_Randriantefy
[pem_pos=0, final_scores=0.69104296,
final_local_scores=0.8452823, pem=0.0, similarity=29.314623]
139: 18964 Madagascar
[pem_pos=0, final_scores=0.6505502,
final_local_scores=0.7366719, pem=-0.30672407,
similarity=27.741486]
140: 20877000 Elena_Makarova
[pem_pos=0, final_scores=0.5135315,
final_local_scores=0.6398613, pem=0.0, similarity=23.18445]
141: 25391 Russia
[pem_pos=0, final_scores=0.85273045,
final_local_scores=0.9046735, pem=-0.20299491,
similarity=32.19089]
142: 3004436 Laurence_Courtois
[pem_pos=0, final_scores=0.85477,
final_local_scores=0.98159426, pem=0.0, similarity=33.382442]
143: 3343 Belgium
[pem_pos=0, final_scores=0.7249069,
final_local_scores=0.7184066, pem=-0.24180195,
similarity=26.843357]
144: 14532 Italy
[pem_pos=0, final_scores=0.77746063,
final_local_scores=0.81436676, pem=-0.20612453,
similarity=29.51298]
145: 37491848 Leander_Paes
[pem_pos=0, final_scores=0.677363,
final_local_scores=0.76943195, pem=0.0, similarity=27.051098]
146: 14533 India
[pem_pos=0, final_scores=0.6064101,
final_local_scores=0.6678481, pem=-0.13622136,

similarity=24.760426]
147: 3930413 Marcos_Ondruska
[pem_pos=0, final_scores=0.73964804,
final_local_scores=0.83986527, pem=0.0, similarity=29.152967]
148: 17416221 South_Africa
[pem_pos=0, final_scores=0.58746964,
final_local_scores=0.6294012, pem=-0.26251382,
similarity=24.299892]
149: 2876758 Jan_Siemerink
[pem_pos=0, final_scores=0.5371706,
final_local_scores=0.62908906, pem=0.0, similarity=22.862986]
150: 21148 Netherlands
[pem_pos=0, final_scores=0.66203886,
final_local_scores=0.66489756, pem=-0.24940795,
similarity=25.287903]
151: 3254714 Carl-Uwe_Steib
[pem_pos=0, final_scores=0.5825657,
final_local_scores=0.6888208, pem=0.0, similarity=24.645498]
152: 11867 Germany
[pem_pos=0, final_scores=0.8981908,
final_local_scores=0.91758734, pem=-0.15529114,
similarity=32.31685]
153: 4026099 Neville_Godwin
[pem_pos=0, final_scores=0.8502336,
final_local_scores=0.9871747, pem=0.0, similarity=33.548973]
154: 17416221 South_Africa
[pem_pos=0, final_scores=0.6255147,
final_local_scores=0.67095965, pem=-0.26251382,
similarity=25.54008]
155: 13833030 Tomás_Carbonell
[pem_pos=0, final_scores=0.61182576,
final_local_scores=0.7279243, pem=0.0, similarity=25.812426]
156: 26667 Spain
[pem_pos=0, final_scores=0.60412496,
final_local_scores=0.6035834, pem=-0.18857051,
similarity=23.127321]
157: 1915197 Jim_Grabb
[pem_pos=0, final_scores=0.64669883,
final_local_scores=0.75650966, pem=0.0, similarity=26.66547]
158: 3434750 United_States
[pem_pos=0, final_scores=0.50246197,
final_local_scores=0.49809057, pem=-0.14075686,
similarity=19.719185]
159: 3823337 Sandon_Stolle
[pem_pos=0, final_scores=0.7639106,
final_local_scores=0.87999636, pem=0.0, similarity=30.350561]
160: 4689264 Australia
[pem_pos=0, final_scores=0.75702804,
final_local_scores=0.79159284, pem=-0.17168477,
similarity=28.64607]
161: 4224688 Alexandra_Fusai
[pem_pos=0, final_scores=0.8409509,

```
final_local_scores=0.9665341, pem=0.0, similarity=32.933018]
162: 5843419 France
    [pem_pos=0, final_scores=0.7151779,
final_local_scores=0.7163819, pem=-0.16240834,
similarity=26.35118]
163: 2442075 Jill_Craybas
    [pem_pos=0, final_scores=0.73184705,
final_local_scores=0.8309701, pem=0.0, similarity=28.887518]
164: 3434750 United_States
    [pem_pos=0, final_scores=0.51979846,
final_local_scores=0.51702803, pem=-0.14075686,
similarity=20.284317]
165: 15573 Japan
    [pem_pos=0, final_scores=0.8147109,
final_local_scores=0.8653367, pem=-0.13549899,
similarity=30.649952]
166: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.38013905,
final_local_scores=0.37667412, pem=-0.6536551,
similarity=18.8851]
167: 1953990 Nathalie_Dechy
    [pem_pos=0, final_scores=0.7065658,
final_local_scores=0.8131382, pem=0.0, similarity=28.35538]
168: 5843419 France
    [pem_pos=0, final_scores=0.75548273,
final_local_scores=0.7604089, pem=-0.16240834,
similarity=27.665031]
169: 11867 Germany
    [pem_pos=0, final_scores=0.84709126,
final_local_scores=0.8617689, pem=-0.15529114,
similarity=30.651115]
170: 3434750 United_States
    [pem_pos=0, final_scores=0.4955355,
final_local_scores=0.49052447, pem=-0.14075686,
similarity=19.4934]
171: 26667 Spain
    [pem_pos=0, final_scores=0.64179426,
final_local_scores=0.6447314, pem=-0.18857051,
similarity=24.355259]
172: 1944281 Els_Callens
    [pem_pos=0, final_scores=0.5982606,
final_local_scores=0.705917, pem=0.0, similarity=25.155684]
173: 3343 Belgium
    [pem_pos=0, final_scores=0.7350311,
final_local_scores=0.7294658, pem=-0.24180195,
similarity=27.173384]
174: 11642302 Nicole_Bradtke
    [pem_pos=0, final_scores=0.94989926,
final_local_scores=1.0924425, pem=0.0, similarity=36.690376]
175: 4689264 Australia
    [pem_pos=0, final_scores=0.73356444,
final_local_scores=0.7659624, pem=-0.17168477,
```

```
similarity=27.881208]
176: 14532 Italy
    [pem_pos=0, final_scores=0.81016237,
final_local_scores=0.8500885, pem=-0.20612453,
similarity=30.578985]
177: 3434750 United_States
    [pem_pos=0, final_scores=0.53957194,
final_local_scores=0.53862756, pem=-0.14075686,
similarity=20.92889]
178: 1514559 Ai_Sugiyama
    [pem_pos=0, final_scores=0.6634075,
final_local_scores=0.7563673, pem=0.0, similarity=26.661222]
179: 15573 Japan
    [pem_pos=0, final_scores=0.77098686,
final_local_scores=0.8175748, pem=-0.13549899,
similarity=29.22464]
180: 11510693 Jana_Kandarr
    [pem_pos=0, final_scores=0.44514984,
final_local_scores=0.5495371, pem=0.0, similarity=20.488997]
181: 11867 Germany
    [pem_pos=0, final_scores=0.85270494,
final_local_scores=0.867901, pem=-0.15529114,
similarity=30.834108]
1: 3523882 National_Tennis_Centre_(United_Kingdom)
    [pem_pos=1, final_scores=0.049232885,
final_local_scores=0.1150257, pem=-0.7567729,
similarity=11.637757]
```

Sample: b'438_BASEBALL'

BASEBALL - [1 1 CUBS 1] EDGE [2 2 BRAVES 2] WITH RUN IN TOP OF NINTH . [3 3 ATLANTA 3] 1996-08-25 [4 4 Brian McRae 4] singled in [5 5 Tyler Houston 5] in the top of the ninth inning to snap a tie as the [6 6 Chicago Cubs 6] avoided a three-game sweep with 3-2 victory over the [7 7 Atlanta Braves 7] on Sunday . The [8 8 Braves 8] scored four runs in the ninth for a 6-5 victory on Saturday . [9 9 Kevin Foster 9] (5-2) won his second straight start , allowing two runs and six hits with two walks and three strikeouts over eight innings . " The biggest thing was my fastball , I was able to rotate it pretty good , " Foster said . " Also , I was able to keep my changeup down . " At [10 10 Colorado 10] , [11 11 Vinny Castilla 11] homered twice and drove in four runs and [12 12 Larry Walker 12] went 3-for-4 with a homer and three [13 13 RBI 13] as the [14 14 Colorado Rockies 14] outslugged the [15 15 Pittsburgh Pirates 15] 13-9 in the rubber game of a three-game series . [16 16 Castilla 16] 's first homer of the game , a solo shot in the seventh off reliever [17 17 Marc Wilkins 17] (3-1) extended [18 18 Colorado 18] 's lead to 9-7 . He added a three-run homer in the eighth off John Ericks to make it 13-8 . At 1 Florida , [19 19 Edgar Renteria 19] 's two-out single in the bottom of the ninth inning scored [20 20 Jesus Tavarez 20] with the winning run as the [2

1 21 Florida Marlins 21] edged the [22 22 Cincinnati Reds 22] 6-5 . " Right after Edgar made contact , I knew I had to score , " said Tavarez . " I knew I would score even if he fielded it cleanly , he could n't throw me out . " " Edgar is a tremendous player right now , " said 2 Florida manager [23 2 3 John Boles 23] . " But I ca n't wait to see how good he 'll be when he grows up . " In [24 24 San Francisco 24] , Osvaldo Fernandez fired a seven-hitter and [25 25 Trenidad Hubbard 25] belted a two-run homer as the [26 26 San Francisco Giants 26] ended a three-game losing streak by defeating the [27 27 Montreal Expos 27] , 7-2 . Fernandez (6-13) allowed two runs , walked one and struck out eight for his second career complete game , both against [28 28 Montreal 28] . In [29 29 Los Angeles 29] , [30 30 Greg Gagne 30] had a run-scoring single and [31 31 Chad Curtis 31] drew a bases-loaded walk in the bottom of the eighth inning as the [32 32 Los Angeles Dodgers 32] rallied for a 6-5 victory and a three-game sweep of the [33 33 New York Mets 33] . " It was one of these games where you get three straight pinch-hits and a walk of a pinch-hitter , that 's how you win pennants , " [34 34 Dodgers 34] manager [35 35 Bill Russell 35] said . " [36 36 Mike Piazza 36]] In [37 37 San Diego 37] , [38 38 Steve Finley 38] and [39 39 Jody Reed 39] drove in three runs apiece as the [40 40 San Diego Padres 40] built a six-run lead after three innings and cruised to an 11-2 victory over the [41 41 Philadelphia Phillies 41] . [42 42 Ken Caminiti 42] added two [43 43 RBI 43] for the [44 44 Padres 44] , who have won six of their last seven games and remained one game ahead of the [45 45 Los Angeles Dodgers 45] in the [46 46 National League West 46] . In [47 47 Houston 47] , [48 48 Jeff Bagwell 48] homered and [49 49 Donne Wall 49] allowed one run over seven innings as the [50 50 Houston Astros 50] defeated the [51 51 St. Louis Cardinals 51] 4-1 . [52 52 Wall 52] (8-4) allowed three hits , walked two and struck out seven as the [53 53 Astros 53] moved 1-1/2 games ahead of the [54 54 Cardinals 54] for the lead in the [55 55 National League Central 55] . He left the game with a knot in his right shoulder .

1: 6654 Chicago_Cubs
2: 2140 Atlanta_Braves
3: 3138 Atlanta
4: 6082884 Brian_McRae
5: 9713764 Tyler_Houston
6: 6654 Chicago_Cubs
7: 2140 Atlanta_Braves
8: 2140 Atlanta_Braves
9: 9713559 Kevin_Foster_(baseball)
10: 6669 Colorado_Rockies
11: 933906 Vinny_Castilla
12: 394237 Larry_Walker
13: 300747 Run_(baseball)
14: 6669 Colorado_Rockies
15: 24270 Pittsburgh_Pirates

```
16: 933906 Vinny_Castilla
17: 12572554 Marc_Wilkins_(baseball)
18: 6669 Colorado_Rockies
19: 849629 Edgar_Rentería
20: 18604283 Jesús_Tavárez
21: 11278 Miami_Marlins
22: 6671 Cincinnati_Reds
23: 3947815 John_Boles_(baseball)
24: 49728 San_Francisco
25: 5107869 Trenidad_Hubbard
26: 28416 San_Francisco_Giants
27: 20153 Montreal_Expos
28: 20153 Montreal_Expos
29: 18110 Los_Angeles
30: 4516364 Greg_Gagne_(baseball)
31: 1821553 Chad_Curtis
32: 18213 Los_Angeles_Dodgers
33: 21728 New_York_Mets
34: 18213 Los_Angeles_Dodgers
35: 3005044 Bill_Russell_(baseball)
36: 182778 Mike_Piazza
37: 28504 San_Diego
38: 1035796 Steve_Finley
39: 2379644 Jody_Reed
40: 28417 San_Diego_Padres
41: 23741 Philadelphia_Phillies
42: 857813 Ken_Caminiti
43: 300747 Run_(baseball)
44: 28417 San_Diego_Padres
45: 18213 Los_Angeles_Dodgers
46: 2456474 National_League_West
47: 13774 Houston
48: 390781 Jeff_Bagwell
49: 6333192 Donne_Wall
50: 13894 Houston_Astros
51: 23408844 St._Louis_Cardinals
52: 6333192 Donne_Wall
53: 13894 Houston_Astros
54: 23408844 St._Louis_Cardinals
55: 2456377 National_League_Central
1: 11278 Miami_Marlins
    [winner: 18933066 Florida, pem_pos=0,
final_scores=-0.4328574, final_local_scores=-0.42138776,
pem=-0.15405534, similarity=-7.6475616|
     RECALL MISS: 11278 Miami_Marlins]
2: 11278 Miami_Marlins
    [winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.6668214, final_local_scores=-0.726341,
pem=-4.620281, similarity=7.5400558|
     RECALL MISS: 11278 Miami_Marlins]
1: 6654 Chicago_Cubs
    [pem_pos=0, final_scores=0.5686522,
```

```
final_local_scores=0.5127129, pem=-0.2883288,
similarity=20.958065]
2: 2140 Atlanta_Braves
    [pem_pos=0, final_scores=0.48408294,
final_local_scores=0.4192096, pem=-0.28448334,
similarity=18.146828]
3: 3138 Atlanta
    [pem_pos=0, final_scores=0.50102746,
final_local_scores=0.56781214, pem=-0.23152597,
similarity=22.293434]
4: 6082884 Brian_McRae
    [pem_pos=0, final_scores=0.5365111,
final_local_scores=0.5498919, pem=0.0, similarity=20.499584]
5: 9713764 Tyler_Houston
    [pem_pos=0, final_scores=0.5429463,
final_local_scores=0.6125168, pem=0.0, similarity=22.368437]
6: 6654 Chicago_Cubs
    [pem_pos=0, final_scores=0.54106766,
final_local_scores=0.4825809, pem=-0.085107334,
similarity=18.953716]
7: 2140 Atlanta_Braves
    [pem_pos=0, final_scores=0.45606196,
final_local_scores=0.38860095, pem=-0.06390974,
similarity=16.033892]
8: 2140 Atlanta_Braves
    [pem_pos=0, final_scores=0.53117985,
final_local_scores=0.47065586, pem=-0.28448334,
similarity=19.682087]
9: 9713559 Kevin_Foster_(baseball)
    [pem_pos=1, final_scores=0.3585921,
final_local_scores=0.43155468, pem=-0.6441395,
similarity=20.471096]
10: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=0.70051676,
final_local_scores=0.65786093, pem=-4.482383,
similarity=48.09749]
11: 933906 Vinny_Castilla
    [pem_pos=0, final_scores=0.75497895,
final_local_scores=0.77888244, pem=0.0, similarity=27.333118]
12: 394237 Larry_Walker
    [pem_pos=0, final_scores=0.75003636,
final_local_scores=0.7811063, pem=0.0, similarity=27.399483]
13: 300747 Run_(baseball)
    [pem_pos=2, final_scores=0.30588305,
final_local_scores=0.2657675, pem=-1.0986123,
similarity=17.99517]
14: 6669 Colorado_Rockies
    [pem_pos=0, final_scores=0.47817147,
final_local_scores=0.4149825, pem=-0.10158596,
similarity=17.026058]
15: 24270 Pittsburgh_Pirates
    [pem_pos=0, final_scores=0.54314846,
```

```
final_local_scores=0.48415112, pem=-0.15187988,
similarity=19.363693]
16: 933906 Vinny_Castilla
    [pem_pos=0, final_scores=1.0519664,
final_local_scores=1.1065445, pem=0.0, similarity=37.11121]
17: 12572554 Marc_Wilkins_(baseball)
    [pem_pos=1, final_scores=0.41165286,
final_local_scores=0.48479193, pem=-0.0139862485,
similarity=18.63293]
18: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=0.8580597,
final_local_scores=0.82995266, pem=-4.482383,
similarity=53.23305]
19: 849629 Edgar_Rentería
    [pem_pos=0, final_scores=0.4805202,
final_local_scores=0.4719091, pem=0.0, similarity=18.172422]
20: 18604283 Jesús_Tavárez
    [pem_pos=0, final_scores=0.47810954,
final_local_scores=0.5493606, pem=0.0, similarity=20.483728]
21: 11278 Miami_Marlins
    [pem_pos=0, final_scores=0.4552366,
final_local_scores=0.39298153, pem=0.0, similarity=15.817065]
22: 6671 Cincinnati_Reds
    [pem_pos=0, final_scores=0.4971581,
final_local_scores=0.4357106, pem=-0.106723964,
similarity=17.672567]
23: 3947815 John_Boles_(baseball)
    [pem_pos=1, final_scores=0.490018,
final_local_scores=0.5154868, pem=-0.81301135,
similarity=23.89415]
24: 49728 San_Francisco
    [pem_pos=0, final_scores=0.5666694,
final_local_scores=0.64362514, pem=-0.09025385,
similarity=23.787586]
25: 5107869 Trenidad_Hubbard
    [pem_pos=0, final_scores=0.5146119,
final_local_scores=0.58967966, pem=0.0, similarity=21.68693]
26: 28416 San_Francisco_Giants
    [pem_pos=0, final_scores=0.5588448,
final_local_scores=0.5009053, pem=-0.07570464,
similarity=19.449419]
27: 20153 Montreal_Expos
    [pem_pos=0, final_scores=0.33923584,
final_local_scores=0.26686722, pem=-0.14979224,
similarity=12.868156]
28: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.63726497,
final_local_scores=0.5924188, pem=-4.5098605,
similarity=46.293995]
29: 18110 Los_Angeles
    [pem_pos=0, final_scores=0.5763003,
final_local_scores=0.6538025, pem=-0.100652285,
```

```
similarity=24.147846]
30: 4516364 Greg_Gagne_(baseball)
    [pem_pos=1, final_scores=0.50090057,
final_local_scores=0.5253721, pem=-0.19170615,
similarity=20.810392]
31: 1821553 Chad_Curtis
    [pem_pos=0, final_scores=0.6119443,
final_local_scores=0.6394592, pem=0.0, similarity=23.172451]
32: 18213 Los_Angeles_Dodgers
    [pem_pos=0, final_scores=0.5554353,
final_local_scores=0.49471647, pem=-0.105150424,
similarity=19.424864]
33: 21728 New_York_Mets
    [pem_pos=0, final_scores=0.4927271,
final_local_scores=0.4305535, pem=-0.065782696,
similarity=17.296024]
34: 18213 Los_Angeles_Dodgers
    [pem_pos=0, final_scores=0.5573509,
final_local_scores=0.49680895, pem=-0.20461318,
similarity=20.028202]
35: 3005044 Bill_Russell_(baseball)
    [pem_pos=1, final_scores=0.4215414,
final_local_scores=0.42515618, pem=-0.8213816,
similarity=21.244022]
36: 182778 Mike_Piazza
    [pem_pos=0, final_scores=0.5664613,
final_local_scores=0.5553509, pem=0.0, similarity=20.66249]
37: 28504 San_Diego
    [pem_pos=0, final_scores=0.57343394,
final_local_scores=0.63973653, pem=-0.18777876,
similarity=24.201897]
38: 1035796 Steve_Finley
    [pem_pos=0, final_scores=0.635985,
final_local_scores=0.6376827, pem=0.0, similarity=23.119436]
39: 2379644 Jody_Reed
    [pem_pos=0, final_scores=0.5163128,
final_local_scores=0.5295542, pem=0.0, similarity=19.892666]
40: 28417 San_Diego_Padres
    [pem_pos=0, final_scores=0.4821567,
final_local_scores=0.4174984, pem=-0.05076447,
similarity=16.824762]
41: 23741 Philadelphia_Phillies
    [pem_pos=0, final_scores=0.45944995,
final_local_scores=0.39343566, pem=-0.09993234,
similarity=16.374065]
42: 857813 Ken_Caminiti
    [pem_pos=0, final_scores=0.65769786,
final_local_scores=0.6854092, pem=0.0, similarity=24.54369]
43: 300747 Run_(baseball)
    [pem_pos=2, final_scores=0.3829419,
final_local_scores=0.34994256, pem=-1.0986123,
similarity=20.507122]
```

```

44: 28417 San_Diego_Padres
    [pem_pos=0, final_scores=0.6168413,
final_local_scores=0.56462085, pem=-0.049586933,
similarity=21.208786]
45: 18213 Los_Angeles_Dodgers
    [pem_pos=0, final_scores=0.59387785,
final_local_scores=0.53670913, pem=-0.105150424,
similarity=20.678009]
46: 2456474 National_League_West
    [pem_pos=0, final_scores=0.362584,
final_local_scores=0.34984502, pem=0.0, similarity=14.529786]
47: 13774 Houston
    [pem_pos=1, final_scores=0.7202377,
final_local_scores=0.7956023, pem=-0.3232624,
similarity=29.590025]
48: 390781 Jeff_Bagwell
    [pem_pos=0, final_scores=0.6236325,
final_local_scores=0.6313665, pem=0.0, similarity=22.930948]
49: 6333192 Donne_Wall
    [pem_pos=0, final_scores=0.7711944,
final_local_scores=0.83022505, pem=0.0, similarity=28.865284]
50: 13894 Houston_Astros
    [pem_pos=0, final_scores=0.41324788,
final_local_scores=0.33941767, pem=-0.049621258,
similarity=14.4884615]
51: 23408844 St._Louis_Cardinals
    [pem_pos=0, final_scores=0.42905378,
final_local_scores=0.36325532, pem=-0.2366796,
similarity=16.217077]
52: 6333192 Donne_Wall
    [pem_pos=0, final_scores=1.3326055,
final_local_scores=1.4478233, pem=0.0, similarity=47.295647]
53: 13894 Houston_Astros
    [pem_pos=0, final_scores=0.5134453,
final_local_scores=0.44886816, pem=-0.294573,
similarity=19.08677]
54: 23408844 St._Louis_Cardinals
    [pem_pos=0, final_scores=0.46526986,
final_local_scores=0.40281594, pem=-1.1265292,
similarity=22.236784]
55: 2456377 National_League_Central
    [pem_pos=0, final_scores=0.44032365,
final_local_scores=0.45203167, pem=0.0, similarity=17.579239]

```

Sample: b'439_BASEBALL'

```

BASEBALL - [1 1 MAJOR LEAGUE 1] STANDINGS AFTER SUNDAY 'S
GAMES . [2 2 NEW YORK 2] 1996-08-25 [3 3 Major League Baseball
3] standings after games played on Sunday ( tabulate under
won , lost , winning percentage and games behind ) : [4 4 AME
RICAN LEAGUE EASTERN DIVISION 4] W L PCT GB [1 1 NEW YORK 1]
74 55 .574 - [5 5 BALTIMORE 5] 68 61 .527 6 [6 6 BOSTON 6] 66
65 .504 9 [7 7 TORONTO 7] 61 70 .466 14 [8 8 DETROIT 8] 47

```

83 .362 27 1/2 [9 9 CENTRAL DIVISION 9] [10 10 CLEVELAND 10]
77 53 .592 - 2 CHICAGO 70 62 .530 8 [11 11 MINNESOTA 11] 65
65 .500 12 [12 12 MILWAUKEE 12] 62 69 .473 15 1/2 [13 13 KANS
AS CITY 13] 59 73 .447 19 [14 14 WESTERN DIVISION 14] [15 15 T
EXAS 15] 75 56 .573 - [16 16 SEATTLE 16] 66 63 .512 8 [17 17 0
AKLND 17] 63 70 .474 13 3 CALIFORNIA 61 69 .469 13 1/2
MONDAY , AUGUST 26TH SCHEDULE [18 18 CLEVELAND 18] AT [19 19 D
ETROIT 19] [20 20 OAKLAND 20] AT [21 21 BALTIMORE 21] [22 22 M
INNESOTA 22] AT [23 23 TORONTO 23] [24 24 MILWAUKEE 24] AT 4 C
HICAGO [25 25 BOSTON 25] AT 5 CALIFORNIA [2 6 NEW YORK 2] AT [
26 26 SEATTLE 26] [27 27 NATIONAL LEAGUE EASTERN DIVISION 27]
W L PCT GB [28 28 ATLANTA 28] 81 48 .628 - [29 29 MONTREAL 29]
70 59 .543 11 7 FLORIDA 61 70 .466 21 [30 30 NEW YORK 30] 59
72 .450 23 [31 31 PHILADELPHIA 31] 53 78 .405 29 [32 32 CENTR
AL DIVISION 32] [33 33 HOUSTON 33] 70 61 .534 - ST LOUIS 68
62 .523 1 1/2 8 CHICAGO 64 64 .500 4 1/2 [34 34 CINCINNATI 34]
64 65 .496 5 [35 35 PITTSBURGH 35] 55 75 .423 14 1/2 [36 36 WE
STERN DIVISION 36] [37 37 SAN DIEGO 37] 72 60 .545 - 9 LOS AN
GELES 70 60 .538 1 [38 38 COLORADO 38] 68 63 .519 3 1/2 [39 39
SAN FRANCISCO 39] 55 73 .430 15 MONDAY , AUGUST 26TH SCHEDULE
[40 40 PHILADELPHIA 40] AT [41 41 SAN FRANCISCO 41] ST LOUIS
AT [42 42 HOUSTON 42] [43 43 CINCINNATI 43] AT [44 44 COLORADO
44]

1: 38776 Major_League_Baseball
2: 645042 New_York_City
3: 38776 Major_League_Baseball
4: 2004582 American_League_East
5: 4340 Baltimore_Orioles
6: 4339 Boston_Red_Sox
7: 30847 Toronto_Blue_Jays
8: 8579 Detroit_Tigers
9: 2456377 National_League_Central
10: 6652 Cleveland_Indians
11: 20050 Minnesota_Twins
12: 20152 Milwaukee_Brewers
13: 16883 Kansas_City_Royals
14: 2456474 National_League_West
15: 30857 Texas_Rangers_(baseball)
16: 27660 Seattle_Mariners
17: 22523 Oakland_Athletics
18: 6652 Cleveland_Indians
19: 8579 Detroit_Tigers
20: 22523 Oakland_Athletics
21: 4340 Baltimore_Orioles
22: 20050 Minnesota_Twins
23: 30847 Toronto_Blue_Jays
24: 20152 Milwaukee_Brewers
25: 4339 Boston_Red_Sox
26: 27660 Seattle_Mariners
27: 2456294 National_League_East
28: 2140 Atlanta_Braves
29: 20153 Montreal_Expos

30: 4848143 New_York_Yankees
31: 23741 Philadelphia_Phillies
32: 2456377 National_League_Central
33: 13894 Houston_Astros
34: 6671 Cincinnati_Reds
35: 24270 Pittsburgh_Pirates
36: 2456474 National_League_West
37: 28417 San_Diego_Padres
38: 6669 Colorado_Rockies
39: 28416 San_Francisco_Giants
40: 23741 Philadelphia_Phillies
41: 28416 San_Francisco_Giants
42: 13894 Houston_Astros
43: 6671 Cincinnati_Reds
44: 6669 Colorado_Rockies
1: 645042 New_York_City
 [winner: 4848143 New_York_Yankees, pem_pos=11,
final_scores=0.2956952, final_local_scores=0.20991808,
pem=-5.0692368, similarity=37.921387]
 gt: 645042 New_York_City, pem_pos=1,
final_scores=0.051055238, final_local_scores=0.0838722,
pem=-1.7906687, similarity=16.33056]
2: 5945 Chicago_White_Sox
 [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.7312331, final_local_scores=-0.82206947,
pem=-4.591283, similarity=4.5256295]
 RECALL MISS: 5945 Chicago_White_Sox]
3: 1360083 Los_Angeles_Angels_ofAnaheim
 [winner: 7590031 California_Golden_Bears_football,
pem_pos=1, final_scores=-0.86890525,
final_local_scores=-0.9543023, pem=-4.4771376,
similarity=-0.04120057]
 RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
4: 5945 Chicago_White_Sox
 [winner: 6886 Chicago, pem_pos=0,
final_scores=-0.6483413, final_local_scores=-0.70828575,
pem=-0.20658566, similarity=-15.923503]
 RECALL MISS: 5945 Chicago_White_Sox]
5: 1360083 Los_Angeles_Angels_ofAnaheim
 [winner: 5407 California, pem_pos=0,
final_scores=-0.7610778, final_local_scores=-0.78981805,
pem=-0.08662026, similarity=-19.008978]
 RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
6: 645042 New_York_City
 [winner: 4848143 New_York_Yankees, pem_pos=11,
final_scores=0.17036308, final_local_scores=0.073011816,
pem=-5.0692368, similarity=33.83583]
 gt: 645042 New_York_City, pem_pos=1,
final_scores=0.07256208, final_local_scores=0.10736519,
pem=-1.7906687, similarity=17.03164]
7: 11278 Miami_Marlins
 [winner: 70551 Florida_Panthers, pem_pos=5,

```
final_scores=-0.7287819, final_local_scores=-0.7982447,
pem=-4.620281, similarity=5.394306|
    RECALL MISS: 11278 Miami_Marlins]
8: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.7255909, final_local_scores=-0.8159062,
pem=-4.591283, similarity=4.709553|
    RECALL MISS: 5945 Chicago_White_Sox]
9: 1360083 Los_Angeles_Angels_of_Anheim
    [winner: 18213 Los_Angeles_Dodgers, pem_pos=15,
final_scores=-0.15750276, final_local_scores=-0.2869594,
pem=-4.7524934, similarity=21.37107|
    RECALL MISS: 1360083 Los_Angeles_Angels_of_Anheim]
1: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.5947402,
final_local_scores=0.547497, pem=-0.47792417,
similarity=23.02714]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.3727724,
final_local_scores=0.43511647, pem=-1.7906687,
similarity=26.812391]
3: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.53948665,
final_local_scores=0.4871409, pem=0.0, similarity=18.626968]
4: 2004582 American_League_East
    [pem_pos=0, final_scores=0.4762742,
final_local_scores=0.4737479, pem=0.0, similarity=18.227295]
5: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.2632324,
final_local_scores=1.2644598, pem=-4.046561,
similarity=63.82954]
6: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.1043936,
final_local_scores=1.094693, pem=-4.418931,
similarity=60.78837]
7: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.73997384,
final_local_scores=0.70241964, pem=-4.422637,
similarity=49.102303]
8: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.2709826,
final_local_scores=1.2733364, pem=-4.1401463,
similarity=64.60337]
9: 2456377 National_League_Central
    [pem_pos=0, final_scores=0.57554317,
final_local_scores=0.59251183, pem=-0.98951703,
similarity=27.152594]
10: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.2680196,
final_local_scores=1.2689784, pem=-4.191953,
similarity=64.75505]
11: 20050 Minnesota_Twins
```

```
[pem_pos=9, final_scores=1.2912078,
final_local_scores=1.2976084, pem=-4.2906566,
similarity=66.146194]
12: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.55244833,
final_local_scores=0.49285334, pem=-3.40205,
similarity=37.298317]
13: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.8177585,
final_local_scores=0.7836708, pem=-3.5915844,
similarity=47.007607]
14: 2456474 National_League_West
    [pem_pos=1, final_scores=0.56962043,
final_local_scores=0.57361805, pem=-1.2952092,
similarity=28.25117]
15: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.3083768,
final_local_scores=1.3187611, pem=-4.928425,
similarity=70.24572]
16: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.7105391,
final_local_scores=0.6713005, pem=-3.8567436,
similarity=45.09623]
17: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.3880098,
final_local_scores=1.4016745, pem=-3.9188845,
similarity=67.22998]
18: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.1435055,
final_local_scores=1.1329656, pem=-4.191953,
similarity=60.696163]
19: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=0.99705356,
final_local_scores=0.9741104, pem=-4.1401463,
similarity=55.67387]
20: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.2973981,
final_local_scores=1.302695, pem=-3.9188845,
similarity=64.27624]
21: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.154662,
final_local_scores=1.1458633, pem=-4.046561,
similarity=60.290382]
22: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=1.2374692,
final_local_scores=1.2389071, pem=-4.2906566,
similarity=64.394424]
23: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.15807201,
final_local_scores=0.06681085, pem=-4.422637,
similarity=30.134468]
24: 20152 Milwaukee_Brewers
```

```
[pem_pos=4, final_scores=0.5332128,
final_local_scores=0.4718414, pem=-3.40205,
similarity=36.67128]
25: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.0771706,
final_local_scores=1.064956, pem=-4.418931,
similarity=59.900955]
26: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.7217541,
final_local_scores=0.68355113, pem=-3.8567436,
similarity=45.46181]
27: 2456294 National_League_East
    [pem_pos=0, final_scores=0.43155545,
final_local_scores=0.42388654, pem=0.0, similarity=16.739332]
28: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=1.1542711,
final_local_scores=1.1454747, pem=-4.136474,
similarity=60.76775]
29: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.6725574,
final_local_scores=0.6261499, pem=-4.5098605,
similarity=47.3006]
30: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.56066966,
final_local_scores=0.49923664, pem=-5.0692368,
similarity=46.55523]
31: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.259844,
final_local_scores=1.2621508, pem=-4.5189333,
similarity=66.32947]
32: 2456377 National_League_Central
    [pem_pos=0, final_scores=0.5688448,
final_local_scores=0.5851949, pem=-0.98951703,
similarity=26.934242]
33: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.3302356,
final_local_scores=1.3397131, pem=-4.650927,
similarity=69.36189]
34: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.1640085,
final_local_scores=1.1572497, pem=-3.771247,
similarity=59.13298]
35: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.315621,
final_local_scores=1.3214293, pem=-4.03702,
similarity=65.47774]
36: 2456474 National_League_West
    [pem_pos=1, final_scores=0.57923335,
final_local_scores=0.5841188, pem=-1.2952092,
similarity=28.564531]
37: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.91874725,
```

```
final_local_scores=0.8952735, pem=-3.9307637,
similarity=52.182564]
38: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.2049544,
final_local_scores=1.2103021, pem=-4.482383,
similarity=64.583435]
39: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.9573994,
final_local_scores=0.9303121, pem=-4.409061,
similarity=55.82924]
40: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.1904241,
final_local_scores=1.1863201, pem=-4.5189333,
similarity=64.06653]
41: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.9044998,
final_local_scores=0.8725273, pem=-4.409061,
similarity=54.104828]
42: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.274334,
final_local_scores=1.2786491, pem=-4.650927,
similarity=67.53962]
43: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.2223437,
final_local_scores=1.2209721, pem=-3.771247,
similarity=61.03458]
44: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=0.9511232,
final_local_scores=0.93302995, pem=-4.482383,
similarity=56.309082]
1: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.2956952,
final_local_scores=0.20991808, pem=-5.0692368,
similarity=37.921387]
2: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.17036308,
final_local_scores=0.073011816, pem=-5.0692368,
similarity=33.83583]

Sample: b'440_BASEBALL'
BASEBALL - [1 1 BONDS 1] ' CONSECUTIVE GAME STREAK ENDS . [2
2 SAN FRANCISCO 2] 1996-08-25 [3 3 San Francisco Giants 3]
All-Star left fielder [4 4 Barry Bonds 4] did not appear in
Sunday 's 7-2 victory over the [5 5 Montreal Expos 5] ,
ending his consecutive games streak . After appearing as a
pinch-hitter in the previous two games , [6 6 Bonds 6] , who
has been battling a hamstring injury , did not see any action
today , ending his streak at 357 consecutive games . It was
the second-longest streak by an active player in the the
majors behind [7 7 Baltimore 7] 's [8 8 Cal Ripken 8] , who
appeared in his major-league record 2,282nd straight game
today , a 13-0 loss to the [9 9 California Angels 9] . [10 10
```

Bonds 10] has been limited to a pinch-hitting role since an MRI Friday showed a mild strain of his left hamstring . [11 11 Bonds 11] came out of Wednesday 's game against [12 12 New York 12] in the ninth inning after suffering a mild hamstring strain . He was back in the starting lineup Thursday night and went 1-for-2 before exiting in the third inning . The 32-year-old [13 13 Bonds 13] is hitting .307 with 35 homers and 107 [14 14 RBI 14] and has been one of the few bright spots for the last-place [15 15 Giants 15] . [16 16 Chicago Cubs 16] outfielder [17 17 Sammy Sosa 17] had the third-longest streak at 304 games , but that ended earlier this week when he suffered a broken bone in his right hand . [18 18 Atlanta Braves 18] first baseman [19 19 Fred McGriff 19] owns the second-longest streak at 295 games .

```
1: 4375 Barry_Bonds
2: 28416 San_Francisco_Giants
3: 28416 San_Francisco_Giants
4: 4375 Barry_Bonds
5: 20153 Montreal_Expos
6: 4375 Barry_Bonds
7: 4340 Baltimore_Orioles
8: 153856 Cal_Ripken,_Jr.
9: 1360083 Los_Angeles_Angels_of_Anheim
10: 4375 Barry_Bonds
11: 4375 Barry_Bonds
12: 4848143 New_York_Yankees
13: 4375 Barry_Bonds
14: 300747 Run_(baseball)
15: 28416 San_Francisco_Giants
16: 6654 Chicago_Cubs
17: 26920 Sammy_Sosa
18: 2140 Atlanta_Braves
19: 205905 Fred_McGriff
1: 4375 Barry_Bonds
    [pem_pos=0, final_scores=0.63601404,
final_local_scores=0.5905148, pem=0.0, similarity=21.711851]
2: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.4368502,
final_local_scores=0.3813309, pem=-4.409061,
similarity=39.446545]
3: 28416 San_Francisco_Giants
    [pem_pos=0, final_scores=0.5911662,
final_local_scores=0.5498977, pem=-0.07570464,
similarity=20.91145]
4: 4375 Barry_Bonds
    [pem_pos=0, final_scores=0.6384154,
final_local_scores=0.5911557, pem=0.0, similarity=21.730978]
5: 20153 Montreal_Expos
    [pem_pos=0, final_scores=0.30056524,
final_local_scores=0.24630183, pem=-0.14979224,
similarity=12.254444]
6: 4375 Barry_Bonds
```

```
[pem_pos=0, final_scores=0.84441286,
final_local_scores=0.81815886, pem=0.0, similarity=28.505205]
7: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=0.90230626,
final_local_scores=0.89645463, pem=-4.046561,
similarity=52.847534]
8: 153856 Cal_Ripken,_Jr.
    [pem_pos=0, final_scores=0.42599565,
final_local_scores=0.40777892, pem=0.0, similarity=16.25865]
9: 1360083 Los_Angeles_Angels_ofAnaheim
    [pem_pos=0, final_scores=0.49136502,
final_local_scores=0.4479779, pem=-0.016987266,
similarity=17.550646]
10: 4375 Barry_Bonds
    [pem_pos=0, final_scores=0.77604085,
final_local_scores=0.7434728, pem=0.0, similarity=26.276424]
11: 4375 Barry_Bonds
    [pem_pos=0, final_scores=0.84968156,
final_local_scores=0.823914, pem=0.0, similarity=28.67695]
12: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.5172704,
final_local_scores=0.46841568, pem=-5.0692368,
similarity=45.635468]
13: 4375 Barry_Bonds
    [pem_pos=0, final_scores=0.6806871,
final_local_scores=0.6393133, pem=0.0, similarity=23.168097]
14: 300747 Run_(baseball)
    [pem_pos=2, final_scores=0.50067836,
final_local_scores=0.4908765, pem=-1.0986123,
similarity=24.712872]
15: 28416 San_Francisco_Giants
    [pem_pos=0, final_scores=0.42123586,
final_local_scores=0.36427462, pem=-0.9105524,
similarity=19.91212]
16: 6654 Chicago_Cubs
    [pem_pos=0, final_scores=0.5956319,
final_local_scores=0.5579582, pem=-0.085107334,
similarity=21.203125]
17: 26920 Sammy_Sosa
    [pem_pos=0, final_scores=0.78087,
final_local_scores=0.7910044, pem=0.0, similarity=27.694862]
18: 2140 Atlanta_Braves
    [pem_pos=0, final_scores=0.48258406,
final_local_scores=0.4384067, pem=-0.06390974,
similarity=17.520195]
19: 205905 Fred_McGriff
    [pem_pos=0, final_scores=0.8152934,
final_local_scores=0.83468175, pem=0.0, similarity=28.998281]
```

Sample: b'441_SOCCER'

SOCCKER - [1 1 JONK 1] RETURNS TO [2 2 DUTCH 2] SQUAD FOR [3 3 BRAZIL 3] FRIENDLY . [4 4 ROTTERDAM 4] 1996-08-26 [5 5 Dutch 5

coach [6 6 Guus Hiddink 6] on Monday recalled midfielder [7 7 Wim Jonk 7] after a 14-month absence for a friendly against [8 8 World Cup 8] holders [9 9 Brazil 9] in [10 10 Amsterdam 1 0] on Sunday . [11 11 Feyenoord 11] midfielder [12 12 Jean-Paul van Gastel 12] was also named to make his debut in the 18-man squad . [13 13 Hiddink 13] did not name a replacement captain for [14 14 Danny Blind 14] , who announced his retirement from international soccer on Sunday . [15 15 Ronald de Boer 15] and [16 16 Dennis Bergkamp 16] are the likely contenders to lead the team . The 35-year-old [17 17 Blind 17] , who won 42 caps for the [18 18 Netherlands 18] , said he wanted to concentrate on playing for his [19 19 Dutch 19] club [20 20 Ajax Amsterdam 20] . [21 21 AC Milan 21] midfielder [22 22 Edgar Davids 22] , who was sent home early from the [23 23 European 23] championship in [24 24 England 24]] after a clash with the coach , was left out of the squad . Squad : Goalkeepers - [25 25 Edwin van der Sar 25] ([26 26 Ajax 26]) , [27 27 Ed de Goey 27] ([28 28 Feyenoord 28]) . Defenders - [29 29 Frank de Boer 29] ([30 30 Ajax 30]) , [31 31 John Veldman 31] ([32 32 Ajax 32]) , [33 33 Jaap Stam 33] ([34 34 PSV 34]) , [35 35 Arthur Numan 35] ([36 36 PSV 36]) , [37 37 Michael Reiziger 37] ([38 38 AC Milan 38]) , [39 39 Johan de Kock 39] ([40 40 Schalke 40] ' 04) . Midfielders - [41 41 Richard Witschge 41] ([42 42 Ajax 42]) , [43 43 Philip Cocu 43] ([44 44 PSV 44]) , [45 45 Wim Jonk 45] ([46 46 PSV 46]) , [47 47 Aron Winter 47] ([48 48 Internazionale 48]) , [49 49 Jean-Paul van Gastel 49] ([50 5 0 Feyenoord 50]) , [51 51 Clarence Seedorf 51] ([52 52 Real Madrid 52]) . Strikers - [53 53 Ronald de Boer 53] ([54 54 Ajax 54]) , [55 55 Gaston Taument 55] ([56 56 Feyenoord 56]) , [57 57 Jordi Cruyff 57] ([58 58 Manchester United 58]) , [59 59 Dennis Bergkamp 59] ([60 60 Arsenal 60]) .

1: 2295002 Wim_Jonk

2: 21148 Netherlands

3: 3383 Brazil

4: 26049 Rotterdam

5: 21148 Netherlands

6: 1345240 Guus_Hiddink

7: 2295002 Wim_Jonk

8: 11370 FIFA_World_Cup

9: 149286 Brazil_national_football_team

10: 844 Amsterdam

11: 47707 Feyenoord

12: 9420542 Jean-Paul_van_Gastel

13: 1345240 Guus_Hiddink

14: 821908 Danny_Blind

15: 1290197 Ronald_de_Boer

16: 45497 Dennis_Bergkamp

17: 821908 Danny_Blind

18: 9647657 Netherlands_national_football_team

19: 21148 Netherlands

20: 2273 AFC_Ajax

21: 18940588 A.C._Milan
22: 252130 Edgar_Davids
23: 9239 Europe
24: 9316 England
25: 685906 Edwin_van_der_Sar
26: 2273 AFC_Ajax
27: 3097068 Ed_de_Goey
28: 47707 Feyenoord
29: 685945 Frank_de_Boer
30: 2273 AFC_Ajax
31: 7363720 John_Veldman
32: 2273 AFC_Ajax
33: 754207 Jaap_Stam
34: 79820 PSV_Eindhoven
35: 1218987 Arthur_Numan
36: 79820 PSV_Eindhoven
37: 756415 Michael_Reiziger
38: 18940588 A.C._Milan
39: 5836737 Johan_de_Kock
40: 410348 FC_Schalke_04
41: 5823946 Richard_Witschge
42: 2273 AFC_Ajax
43: 1295837 Phillip_Cocu
44: 79820 PSV_Eindhoven
45: 2295002 Wim_Jonk
46: 79820 PSV_Eindhoven
47: 1298502 Aron_Winter
48: 15116 Inter_Milan
49: 9420542 Jean-Paul_van_Gastel
50: 47707 Feyenoord
51: 616584 Clarence_Seedorf
52: 26413 Real_Madrid_C.F.
53: 1290197 Ronald_de_Boer
54: 2273 AFC_Ajax
55: 5824381 Gaston_Taument
56: 47707 Feyenoord
57: 1771982 Jordi_Cruyff
58: 19961 Manchester_United_F.C.
59: 45497 Dennis_Bergkamp
60: 2174 Arsenal_F.C.
1: 2295002 Wim_Jonk
[pem_pos=0, final_scores=0.596346,
final_local_scores=0.6310793, pem=0.0, similarity=22.922379]
2: 21148 Netherlands
[pem_pos=0, final_scores=0.5576351,
final_local_scores=0.5945313, pem=-0.7943452,
similarity=26.151485]
3: 3383 Brazil
[pem_pos=0, final_scores=0.5056923,
final_local_scores=0.60615367, pem=-0.21313342,
similarity=23.337599]
4: 26049 Rotterdam

```
[pem_pos=0, final_scores=0.5148975,
final_local_scores=0.5551749, pem=-0.145035,
similarity=21.445961]
5: 21148 Netherlands
    [pem_pos=0, final_scores=0.39525938,
final_local_scores=0.41716045, pem=-0.7943452,
similarity=20.858385]
6: 1345240 Guus_Hiddink
    [pem_pos=0, final_scores=0.74457234,
final_local_scores=0.7543102, pem=0.0, similarity=26.599834]
7: 2295002 Wim_Jonk
    [pem_pos=0, final_scores=0.713426,
final_local_scores=0.7589714, pem=0.0, similarity=26.738934]
8: 11370 FIFA_World_Cup
    [pem_pos=0, final_scores=0.2282447,
final_local_scores=0.20390803, pem=-1.001087,
similarity=15.618803]
9: 149286 Brazil_national_football_team
    [pem_pos=1, final_scores=0.5236048,
final_local_scores=0.50766, pem=-2.5928838,
similarity=33.339806]
10: 844 Amsterdam
    [pem_pos=0, final_scores=0.5688023,
final_local_scores=0.612802, pem=-0.0642803,
similarity=22.726515]
11: 47707 Feyenoord
    [pem_pos=0, final_scores=0.46367162,
final_local_scores=0.40944415, pem=0.0, similarity=16.308344]
12: 9420542 Jean-Paul_van_Gastel
    [pem_pos=0, final_scores=0.5721816,
final_local_scores=0.6539949, pem=0.0, similarity=23.606226]
13: 1345240 Guus_Hiddink
    [pem_pos=0, final_scores=0.64186376,
final_local_scores=0.64211667, pem=0.0, similarity=23.251755]
14: 821908 Danny_Blind
    [pem_pos=0, final_scores=0.7940653,
final_local_scores=0.82355016, pem=0.0, similarity=28.666092]
15: 1290197 Ronald_de_Boer
    [pem_pos=0, final_scores=0.54814976,
final_local_scores=0.5747704, pem=0.0, similarity=21.242008]
16: 45497 Dennis_Bergkamp
    [pem_pos=0, final_scores=0.7232576,
final_local_scores=0.7234998, pem=0.0, similarity=25.68039]
17: 821908 Danny_Blind
    [pem_pos=0, final_scores=1.0137466,
final_local_scores=1.0666654, pem=0.0, similarity=35.921135]
18: 9647657 Netherlands_national_football_team
    [pem_pos=1, final_scores=0.62484187,
final_local_scores=0.594898, pem=-2.660505,
similarity=36.310894]
19: 21148 Netherlands
    [pem_pos=0, final_scores=0.52542824,
```

```
final_local_scores=0.55935013, pem=-0.7943452,  
similarity=25.10161]  
20: 2273 AFC_Ajax  
    [pem_pos=0, final_scores=0.63772815,  
final_local_scores=0.5916331, pem=0.0, similarity=21.745224]  
21: 18940588 A.C._Milan  
    [pem_pos=0, final_scores=0.6840559,  
final_local_scores=0.6707051, pem=0.0, similarity=24.10489]  
22: 252130 Edgar_Davids  
    [pem_pos=0, final_scores=0.7075407,  
final_local_scores=0.7090618, pem=0.0, similarity=25.24953]  
23: 9239 Europe  
    [pem_pos=0, final_scores=0.46690714,  
final_local_scores=0.57997555, pem=-0.25557294,  
similarity=22.787186]  
24: 9316 England  
    [pem_pos=0, final_scores=0.41842526,  
final_local_scores=0.51496196, pem=-0.19053856,  
similarity=20.493383]  
25: 685906 Edwin_van_der_Sar  
    [pem_pos=0, final_scores=0.73870414,  
final_local_scores=0.7677976, pem=0.0, similarity=27.002325]  
26: 2273 AFC_Ajax  
    [pem_pos=1, final_scores=0.5227271,  
final_local_scores=0.46601194, pem=-1.4356393,  
similarity=25.803669]  
27: 3097068 Ed_de_Goey  
    [pem_pos=0, final_scores=0.5513196,  
final_local_scores=0.6000797, pem=0.0, similarity=21.997288]  
28: 47707 Feyenoord  
    [pem_pos=0, final_scores=0.50301856,  
final_local_scores=0.4524247, pem=0.0, similarity=17.59097]  
29: 685945 Frank_de_Boer  
    [pem_pos=0, final_scores=0.8080271,  
final_local_scores=0.8248742, pem=0.0, similarity=28.705605]  
30: 2273 AFC_Ajax  
    [pem_pos=1, final_scores=0.48496252,  
final_local_scores=0.42475992, pem=-1.4356393,  
similarity=24.572624]  
31: 7363720 John_Veldman  
    [pem_pos=0, final_scores=0.6965386,  
final_local_scores=0.7921857, pem=0.0, similarity=27.730116]  
32: 2273 AFC_Ajax  
    [pem_pos=1, final_scores=0.49532062,  
final_local_scores=0.43607455, pem=-1.4356393,  
similarity=24.910276]  
33: 754207 Jaap_Stam  
    [pem_pos=0, final_scores=0.83739716,  
final_local_scores=0.85951984, pem=0.0, similarity=29.7395]  
34: 79820 PSV_Eindhoven  
    [pem_pos=0, final_scores=0.5590187,  
final_local_scores=0.5107376, pem=0.0, similarity=19.331142]
```

35: 1218987 Arthur_Numan
[pem_pos=0, final_scores=0.827503,
final_local_scores=0.8621012, pem=0.0, similarity=29.816532]
36: 79820 PSV_Eindhoven
[pem_pos=0, final_scores=0.5740779,
final_local_scores=0.5271875, pem=0.0, similarity=19.82204]
37: 756415 Michael_Reiziger
[pem_pos=0, final_scores=0.8226183,
final_local_scores=0.8734904, pem=0.0, similarity=30.15641]
38: 18940588 A.C._Milan
[pem_pos=0, final_scores=0.53764147,
final_local_scores=0.51076955, pem=0.0, similarity=19.332094]
39: 5836737 Johan_de_Kock
[pem_pos=0, final_scores=0.4670347,
final_local_scores=0.5431285, pem=0.0, similarity=20.29775]
40: 410348 FC_Schalke_04
[pem_pos=0, final_scores=0.37304348,
final_local_scores=0.33525962, pem=0.0, similarity=14.094528]
41: 5823946 Richard_Witschge
[pem_pos=0, final_scores=0.9385393,
final_local_scores=0.9935755, pem=0.0, similarity=33.739986]
42: 2273 AFC_Ajax
[pem_pos=1, final_scores=0.505696,
final_local_scores=0.44740802, pem=-1.4356393,
similarity=25.248491]
43: 1295837 Phillip_Cocu
[pem_pos=0, final_scores=0.9128434,
final_local_scores=0.93479174, pem=0.0, similarity=31.985762]
44: 79820 PSV_Eindhoven
[pem_pos=0, final_scores=0.6050376,
final_local_scores=0.5610063, pem=0.0, similarity=20.83126]
45: 2295002 Wim_Jonk
[pem_pos=0, final_scores=0.7968028,
final_local_scores=0.8500479, pem=0.0, similarity=29.456837]
46: 79820 PSV_Eindhoven
[pem_pos=0, final_scores=0.6491043,
final_local_scores=0.6091425, pem=0.0, similarity=22.26774]
47: 1298502 Aron_Winter
[pem_pos=0, final_scores=0.8699408,
final_local_scores=0.91739506, pem=0.0, similarity=31.466612]
48: 15116 Inter_Milan
[pem_pos=0, final_scores=0.3355109,
final_local_scores=0.29591042, pem=0.0, similarity=12.92027]
49: 9420542 Jean-Paul_van_Gastel
[pem_pos=0, final_scores=0.6399926,
final_local_scores=0.7280681, pem=0.0, similarity=25.816717]
50: 47707 Feyenoord
[pem_pos=0, final_scores=0.53247327,
final_local_scores=0.48459947, pem=0.0, similarity=18.551128]
51: 616584 Clarence_Seedorf
[pem_pos=0, final_scores=0.9113992,
final_local_scores=0.95410746, pem=0.0, similarity=32.562183]

```
52: 26413 Real_Madrid_C.F.  
    [pem_pos=0, final_scores=0.36408603,  
final_local_scores=0.32589746, pem=0.0, similarity=13.815143]  
53: 1290197 Ronald_de_Boer  
    [pem_pos=0, final_scores=0.53149897,  
final_local_scores=0.55658185, pem=0.0, similarity=20.699226]  
54: 2273 AFC_Ajax  
    [pem_pos=1, final_scores=0.52594995,  
final_local_scores=0.46953243, pem=-1.4356393,  
similarity=25.908726]  
55: 5824381 Gaston_Taument  
    [pem_pos=0, final_scores=0.76655596,  
final_local_scores=0.8486363, pem=0.0, similarity=29.414713]  
56: 47707 Feyenoord  
    [pem_pos=0, final_scores=0.6096159,  
final_local_scores=0.56886613, pem=0.0, similarity=21.065813]  
57: 1771982 Jordi_Cruyff  
    [pem_pos=0, final_scores=0.9020781,  
final_local_scores=0.9405418, pem=0.0, similarity=32.157356]  
58: 19961 Manchester_United_F.C.  
    [pem_pos=0, final_scores=0.40684742,  
final_local_scores=0.37346363, pem=0.0, similarity=15.234612]  
59: 45497 Dennis_Bergkamp  
    [pem_pos=0, final_scores=0.66686773,  
final_local_scores=0.6619024, pem=0.0, similarity=23.842201]  
60: 2174 Arsenal_F.C.  
    [pem_pos=0, final_scores=0.48356175,  
final_local_scores=0.45711285, pem=-0.15820743,  
similarity=18.591228]
```

Sample: b'442_SOCCER'

```
SOCCER - [1 1 BARCELONA 1] BEAT [1 ATLETICO 1] 5-2 IN [2 2 SU  
PERCUP 2] . [3 3 BARCELONA 3] 1996-08-26 [4 4 Barcelona 4]  
beat [5 5 Atletico Madrid 5] 5-2 ( halftime 2-1 ) in the [6 6  
Spanish Supercup 6] on Sunday : Scorers : [7 7 Barcelona 7] -  
[8 8 Ronaldo 8] ( 5th and 89th minutes ) , [9 9 Giovanni 9]  
( 31st ) , [10 10 Pizzi 10] ( 73rd ) , De la Pena ( 75th ) [11  
11 Atletico Madrid 11] - Esnaider ( 37th ) , Pantic ( 57th ,  
penalty ) Attendance 30,000
```

```
1: 68187 FC_Barcelona  
2: 876551 Supercopa_de_España  
3: 4443 Barcelona  
4: 68187 FC_Barcelona  
5: 84197 Atlético_Madrid  
6: 876551 Supercopa_de_España  
7: 68187 FC_Barcelona  
8: 66164 Ronaldo  
9: 2647669 Giovanni_Silva_de_Oliveira  
10: 2706494 Juan_Antonio_Pizzi  
11: 84197 Atlético_Madrid  
1: 68187 FC_Barcelona  
    [pem_pos=1, final_scores=0.6789637,
```

```
final_local_scores=0.63176364, pem=-1.7089969,
similarity=32.23659]
2: 876551 Supercopa_de_España
    [pem_pos=5, final_scores=0.17772828,
final_local_scores=0.13174736, pem=-2.0589352,
similarity=19.218126]
3: 4443 Barcelona
    [pem_pos=0, final_scores=0.3503595,
final_local_scores=0.38975102, pem=-0.37077636,
similarity=17.737]
4: 68187 FC_Barcelona
    [pem_pos=1, final_scores=0.5128911,
final_local_scores=0.45035452, pem=-1.7089969,
similarity=26.822979]
5: 84197 Atlético_Madrid
    [pem_pos=0, final_scores=0.52907646,
final_local_scores=0.47658718, pem=0.0, similarity=18.312025]
6: 876551 Supercopa_de_España
    [pem_pos=0, final_scores=0.35430646,
final_local_scores=0.32734725, pem=0.0, similarity=13.858408]
7: 68187 FC_Barcelona
    [pem_pos=1, final_scores=0.5800635,
final_local_scores=0.5237302, pem=-1.7089969,
similarity=29.012657]
8: 66164 Ronaldo
    [pem_pos=0, final_scores=0.31007344,
final_local_scores=0.30229273, pem=-0.10865249,
similarity=13.7016]
9: 2647669 Giovanni_Silva_de_Oliveira
    [pem_pos=0, final_scores=0.49586678,
final_local_scores=0.5706309, pem=-1.3095523,
similarity=28.240026]
10: 2706494 Juan_Antonio_Pizzi
    [pem_pos=0, final_scores=0.38376862,
final_local_scores=0.40931445, pem=-0.15201622,
similarity=17.13116]
11: 84197 Atlético_Madrid
    [pem_pos=0, final_scores=0.5179876,
final_local_scores=0.46447432, pem=0.0, similarity=17.950554]
1: 84197 Atlético_Madrid
    [pem_pos=1, final_scores=0.057532236,
final_local_scores=-0.043283343, pem=-2.5207162,
similarity=16.506098]
```

Sample: b'443_SOCCER'
SOCCER - [1 1 AUSTRIA 1] FIRST DIVISION RESULTS /
STANDINGS . [2 2 VIENNA 2] 1996-08-26 Result of an [3 3 Austrian 3] first division soccer match played on Sunday : [4 4 SV Ried 4] 0 [5 5 SV Salzburg 5] 4 Standings (tabulated under played , won , drawn , lost , goals for , against , points) : [6 6 FC Tirol Innsbruck 6] 6 4 2 0 13 5 14 [7 7 SV Salzburg 7] 6 4 2 0 8 1 14 [8 8 Austria Vienna 8] 6 4 2 0 9 5

```
14 [9 9 Sturm Graz 9] 6 2 3 1 8 5 9 1 GAK 6 1 3 2 8 10 6 [10 1
0 Rapid Wien 10] 5 0 5 0 3 3 5 [11 11 SV Ried 11] 6 1 1 4 6 9
4 [12 12 Linzer ASK 12] 5 0 3 2 4 8 3 Admira / Wacker 6 0 3 3
5 10 3 [13 13 FC Linz 13] 6 0 2 4 1 9 2
1: 26964606 Austria
2: 55866 Vienna
3: 26964606 Austria
4: 3004699 SV_Ried
5: 2163477 FC_Red_Bull_Salzburg
6: 5599205 FC_Tirol_Innsbruck
7: 2163477 FC_Red_Bull_Salzburg
8: 895568 FK_Austria_Wien
9: 1595328 SK_Sturm_Graz
10: 656932 SK_Rapid_Wien
11: 3004699 SV_Ried
12: 5644175 LASK_Linz
13: 5644199 FC_Linz
1: 887879 Grazer_AK
    [winner: 3968641 Nickelodeon_Toys, pem_pos=0,
final_scores=-0.37345263, final_local_scores=-0.21877511,
pem=0.0, similarity=-2.4389727|
     RECALL MISS: 887879 Grazer_AK]
1: 26964606 Austria
    [pem_pos=0, final_scores=0.6236951,
final_local_scores=0.67138976, pem=-0.25262418,
similarity=25.499132]
2: 55866 Vienna
    [pem_pos=0, final_scores=0.5072377,
final_local_scores=0.5626711, pem=-0.12517071,
similarity=21.561638]
3: 26964606 Austria
    [pem_pos=0, final_scores=0.6763763,
final_local_scores=0.72893596, pem=-0.40674555,
similarity=28.054562]
4: 3004699 SV_Ried
    [pem_pos=0, final_scores=0.5247199,
final_local_scores=0.51088256, pem=0.0, similarity=19.335468]
5: 2163477 FC_Red_Bull_Salzburg
    [pem_pos=0, final_scores=0.43089336,
final_local_scores=0.38820773, pem=0.0, similarity=15.674606]
6: 5599205 FC_Tirol_Innsbruck
    [pem_pos=0, final_scores=0.42120355,
final_local_scores=0.43863928, pem=0.0, similarity=17.179585]
7: 2163477 FC_Red_Bull_Salzburg
    [pem_pos=0, final_scores=0.4161808,
final_local_scores=0.37213653, pem=0.0, similarity=15.195009]
8: 895568 FK_Austria_Wien
    [pem_pos=0, final_scores=0.552435,
final_local_scores=0.5143203, pem=0.0, similarity=19.438057]
9: 1595328 SK_Sturm_Graz
    [pem_pos=0, final_scores=0.44043952,
final_local_scores=0.4051003, pem=0.0, similarity=16.178713]
```

```
10: 656932 SK_Rapid_Wien
    [pem_pos=0, final_scores=0.39279604,
final_local_scores=0.3490695, pem=0.0, similarity=14.506643]
11: 3004699 SV_Ried
    [pem_pos=0, final_scores=0.5242262,
final_local_scores=0.51034325, pem=0.0, similarity=19.319374]
12: 5644175 LASK_Linz
    [pem_pos=0, final_scores=0.56595653,
final_local_scores=0.5434873, pem=0.0, similarity=20.308458]
13: 5644199 FC_Linz
    [pem_pos=1, final_scores=0.3182026,
final_local_scores=0.34845066, pem=-0.50743,
similarity=17.24766]
```

Sample: b'444_CRICKET'

CRICKET - [1 1 AUSTRALIA 1] BEAT [2 2 ZIMBABWE 2] BY 125 RUNS IN ONE-DAY MATCH . [3 3 COLOMBO 3] 1996-08-26 [4 4 Australia 4] beat [5 5 Zimbabwe 5] by 125 runs in the first match of the [6 6 Singer World Series 6] one-day (50 overs) cricket tournament on Monday . Scores : [7 7 Australia 7] 263-7 in 50 overs , [8 8 Zimbabwe 8] 138 all out in 41 overs .

```
1: 374833 Australia_national_cricket_team
2: 759317 Zimbabwe_national_cricket_team
3: 56636 Colombo
4: 374833 Australia_national_cricket_team
5: 759317 Zimbabwe_national_cricket_team
6: 28357832 Singer_World_Series
7: 374833 Australia_national_cricket_team
8: 759317 Zimbabwe_national_cricket_team
1: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.6989066,
final_local_scores=0.65435183, pem=-3.4989398,
similarity=42.644657]
2: 759317 Zimbabwe_national_cricket_team
    [pem_pos=1, final_scores=0.93979126,
final_local_scores=0.9204611, pem=-2.2776897,
similarity=43.94454]
3: 56636 Colombo
    [pem_pos=0, final_scores=0.518444,
final_local_scores=0.5793923, pem=-0.19393882,
similarity=22.434605]
4: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.8051587,
final_local_scores=0.77041614, pem=-3.4989398,
similarity=46.108246]
5: 759317 Zimbabwe_national_cricket_team
    [pem_pos=1, final_scores=0.9583214,
final_local_scores=0.9407025, pem=-2.2776897,
similarity=44.54858]
6: 28357832 Singer_World_Series
    [pem_pos=0, final_scores=0.6087159,
final_local_scores=0.6822816, pem=0.0, similarity=24.450357]
```

```
7: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.8683601,
final_local_scores=0.83945405, pem=-3.4989398,
similarity=48.168476]
8: 759317 Zimbabwe_national_cricket_team
    [pem_pos=1, final_scores=0.97277325,
final_local_scores=0.9564889, pem=-2.2776897,
similarity=45.01968]
```

Sample: b'445_CRICKET'

CRICKET - SCOREBOARD-AUSTRALIA V [1 1 ZIMBABWE 1] ONE-DAY SCOREBOARD . [2 2 COLOMBO 2] 1996-08-26 Scoreboard in the [3 3 Singer World Series 3] one-day (50 overs) cricket match between [4 4 Australia 4] and [5 5 Zimbabwe 5] on Monday : [6 6 Australia 6] M. Slater c P. Strang b [7 7 Whittall 7] 50 M. Waugh b P. Strang 18 R. Ponting c and b [8 8 Whittall 8] 53 [9 9 S. Waugh 9] c 1 Campbell b [10 10 Whittall 10] 82 S. Law b [11 11 Streak 11] 20 M. Bevan c 2 Campbell b [12 12 Brandes 12] 9 I. Healy b [13 13 Brandes 13] 5 B. Hogg not out 11 Extras (b-1 lb-8 w-3 nb-3) 15 Total (for seven wickets - 50 overs) 263 Fall of wickets : 1-48 2-92 3-167 4-230 5-240 6-242 7-263 Did not bat : P. Reiffel , D. Flemming , G. McGrath Bowling : [14 14 Streak 14] 10-1-50-1 (2w , 2nb) , [15 15 Brandes 15] 10-1-47-2 (1w) , P. Strang 9-0-41-1 , [16 16 Flower 16] 6-0-28-0 , [17 17 Whittall 17] 10-0-53-3 (1nb) , Decker 3-0-17-0 , [18 18 Shah 18] 2-0-18-0 [19 19 Zimbabwe 19] A. Shah c M. Waugh b [20 20 Hogg 20] 41 G. Flower c [21 21 Ponting 21] b Flemming 7 A. Flower lbw b Flemming 0 A. Campbell lbw b [22 22 McGrath 22] 9 C. Wishart c [23 23 Healy 23] b [24 24 Reiffel 24] 0 G. Whittall b [25 25 Reiffel 25] 11 C. Evans c [26 26 Healy 26] b [27 27 S. Waugh 27] 15 M. Dekker not out 8 P. Strang b M. Waugh 9 H. Streak b M. Waugh 0 E. Brandes c [28 28 Hogg 28] b M. Waugh 17 Extras (lb-4 w-10 nb-7) 21 Total (all out - 41 overs) 138 Fall of wickets : 1-16 2-16 3-33 4-35 5-56 6-98 7-100 8-120 9-120 Bowling : [29 29 McGrath 29] 7-2-13-1 (2w) , Flemming 7-0-24-2 (3w , 3nb) , [30 30 Reiffel 30] 6-1-23-2 (2nb) , [31 31 S. Waugh 31] 7-2-24-1 (1nb , 2w) , [32 32 Hogg 32] 9-2-26-1 (1nb , 3w) , M. Waugh 5-1-24-3 Result : [33 33 Australia 33] won by 125 runs .

```
1: 759317 Zimbabwe_national_cricket_team
2: 56636 Colombo
3: 28357832 Singer_World_Series
4: 374833 Australia_national_cricket_team
5: 759317 Zimbabwe_national_cricket_team
6: 374833 Australia_national_cricket_team
7: 4036552 Andy_Whittall
8: 4036552 Andy_Whittall
9: 156123 Steve_Waugh
10: 4036552 Andy_Whittall
11: 11900310 Denis_Streak
12: 3966081 Eddo_Brandes
```

```
13: 3966081 Eddo_Brandes
14: 897195 Heath_Streak
15: 3966081 Eddo_Brandes
16: 733585 Andy_Flower
17: 4036552 Andy_Whittall
18: 4036464 Ali_Shah
19: 759317 Zimbabwe_national_cricket_team
20: 1384007 Brad_Hogg
21: 462001 Ricky_Ponting
22: 423488 Glenn_McGrath
23: 928267 Ian_Healy
24: 2383377 Paul_Reiffel
25: 2383377 Paul_Reiffel
26: 928267 Ian_Healy
27: 156123 Steve_Waugh
28: 1384007 Brad_Hogg
29: 423488 Glenn_McGrath
30: 2383377 Paul_Reiffel
31: 156123 Steve_Waugh
32: 1384007 Brad_Hogg
33: 374833 Australia_national_cricket_team
1: 2156908 Alistair_Campbell_(cricketer)
    [winner: 35591326 Craig_Campbell_(tennis), pem_pos=0,
final_scores=-0.43110573, final_local_scores=-0.33078593,
pem=0.0, similarity=-5.7815986]
    RECALL MISS: 2156908 Alistair_Campbell_(cricketer)]
2: 2156908 Alistair_Campbell_(cricketer)
    [winner: 35591326 Craig_Campbell_(tennis), pem_pos=0,
final_scores=-0.42264313, final_local_scores=-0.3215418,
pem=0.0, similarity=-5.505736]
    RECALL MISS: 2156908 Alistair_Campbell_(cricketer)]
1: 759317 Zimbabwe_national_cricket_team
    [pem_pos=1, final_scores=0.89164287,
final_local_scores=0.8991551, pem=-2.2776897,
similarity=43.308723]
2: 56636 Colombo
    [pem_pos=0, final_scores=0.519729,
final_local_scores=0.61283, pem=-0.19393882,
similarity=23.432451]
3: 28357832 Singer_World_Series
    [pem_pos=0, final_scores=0.6405562,
final_local_scores=0.71982735, pem=0.0, similarity=25.570797]
4: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.8168233,
final_local_scores=0.8049588, pem=-3.4989398,
similarity=47.13907]
5: 759317 Zimbabwe_national_cricket_team
    [pem_pos=1, final_scores=0.914272,
final_local_scores=0.9238741, pem=-2.2776897,
similarity=44.046387]
6: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.85404783,
```

```
final_local_scores=0.845621, pem=-3.4989398,
similarity=48.35251]
7: 4036552 Andy_Whittall
    [pem_pos=3, final_scores=0.54259783,
final_local_scores=0.63297695, pem=-1.3862944,
similarity=30.51789]
8: 4036552 Andy_Whittall
    [pem_pos=3, final_scores=0.56304246,
final_local_scores=0.6553096, pem=-1.3862944,
similarity=31.184343]
9: 156123 Steve_Waugh
    [pem_pos=0, final_scores=0.57475346,
final_local_scores=0.5536769, pem=0.0, similarity=20.612535]
10: 4036552 Andy_Whittall
    [pem_pos=3, final_scores=0.5858595,
final_local_scores=0.6802338, pem=-1.3862944,
similarity=31.92813]
11: 11900310 Denis_Streak
    [pem_pos=3, final_scores=0.3732345,
final_local_scores=0.46824747, pem=-1.7917595,
similarity=27.807018]
12: 3966081 Eddo_Brandes
    [pem_pos=0, final_scores=0.56038237,
final_local_scores=0.61823225, pem=-0.36647612,
similarity=24.53195]
13: 3966081 Eddo_Brandes
    [pem_pos=0, final_scores=0.5719959,
final_local_scores=0.63091826, pem=-0.36647612,
similarity=24.910526]
14: 897195 Heath_Streak
    [pem_pos=1, final_scores=0.6679521,
final_local_scores=0.68969005, pem=-1.6958603,
similarity=33.89379]
15: 3966081 Eddo_Brandes
    [pem_pos=0, final_scores=0.5485006,
final_local_scores=0.60525316, pem=-0.36647612,
similarity=24.144629]
16: 733585 Andy_Flower
    [pem_pos=14, final_scores=0.7581075,
final_local_scores=0.76425743, pem=-3.7938714,
similarity=47.528343]
17: 4036552 Andy_Whittall
    [pem_pos=3, final_scores=0.6329796,
final_local_scores=0.73170525, pem=-1.3862944,
similarity=33.46414]
18: 4036464 Ali_Shah
    [pem_pos=11, final_scores=0.50876856,
final_local_scores=0.60919416, pem=-4.937935,
similarity=49.12254]
19: 759317 Zimbabwe_national_cricket_team
    [pem_pos=1, final_scores=0.83887535,
final_local_scores=0.84151465, pem=-2.2776897,
```

```
similarity=41.58862]
20: 1384007 Brad_Hogg
    [pem_pos=9, final_scores=0.67212373,
final_local_scores=0.6690971, pem=-3.117661,
similarity=41.01123]
21: 462001 Ricky_Ponting
    [pem_pos=0, final_scores=0.56833917,
final_local_scores=0.54596096, pem=-0.24151042,
similarity=21.695648]
22: 423488 Glenn_McGrath
    [pem_pos=3, final_scores=0.7757705,
final_local_scores=0.7684644, pem=-2.1900785,
similarity=38.932213]
23: 928267 Ian_Healy
    [pem_pos=9, final_scores=0.62942606,
final_local_scores=0.6234777, pem=-3.6584673,
similarity=42.590847]
24: 2383377 Paul_Reiffel
    [pem_pos=0, final_scores=0.6817245,
final_local_scores=0.70372105, pem=0.0, similarity=25.090153]
25: 2383377 Paul_Reiffel
    [pem_pos=0, final_scores=0.5418176,
final_local_scores=0.5508941, pem=0.0, similarity=20.529491]
26: 928267 Ian_Healy
    [pem_pos=9, final_scores=0.65421206,
final_local_scores=0.65055275, pem=-3.6584673,
similarity=43.39882]
27: 156123 Steve_Waugh
    [pem_pos=0, final_scores=0.5096411,
final_local_scores=0.48255152, pem=0.0, similarity=18.490013]
28: 1384007 Brad_Hogg
    [pem_pos=9, final_scores=0.5882657,
final_local_scores=0.5774949, pem=-3.117661,
similarity=38.277637]
29: 423488 Glenn_McGrath
    [pem_pos=3, final_scores=0.8562016,
final_local_scores=0.8563232, pem=-2.1900785,
similarity=41.554092]
30: 2383377 Paul_Reiffel
    [pem_pos=0, final_scores=0.6955901,
final_local_scores=0.7188671, pem=0.0, similarity=25.542141]
31: 156123 Steve_Waugh
    [pem_pos=0, final_scores=0.46982914,
final_local_scores=0.43906295, pem=0.0, similarity=17.192226]
32: 1384007 Brad_Hogg
    [pem_pos=9, final_scores=0.7292909,
final_local_scores=0.7315436, pem=-3.117661,
similarity=42.87476]
33: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.75338894,
final_local_scores=0.7356663, pem=-3.4989398,
similarity=45.07124]
```

Sample: b'446_CRICKET'

CRICKET - [1 1 AUSTRALIA 1] 263-7 IN 50 OVERS V [2 2 ZIMBABWE 2] . [3 3 COLOMBO 3] 1996-08-26 [4 4 Australia 4] scored 263-7 in their 50 overs against [5 5 Zimbabwe 5] in the first day-night limited overs match of the [6 6 Singer World Series 6] tournament on Monday .

1: 374833 Australia_national_cricket_team
2: 759317 Zimbabwe_national_cricket_team
3: 56636 Colombo
4: 374833 Australia_national_cricket_team
5: 759317 Zimbabwe_national_cricket_team
6: 28357832 Singer_World_Series

1: 374833 Australia_national_cricket_team
[pem_pos=1, final_scores=0.7028952,
final_local_scores=0.6661997, pem=-3.4989398,
similarity=42.99822]
2: 759317 Zimbabwe_national_cricket_team
[pem_pos=1, final_scores=0.72960943,
final_local_scores=0.6998895, pem=-2.2776897,
similarity=37.36224]
3: 56636 Colombo
[pem_pos=0, final_scores=0.52282745,
final_local_scores=0.5857238, pem=-0.19393882,
similarity=22.62355]
4: 374833 Australia_national_cricket_team
[pem_pos=1, final_scores=0.80108505,
final_local_scores=0.77345717, pem=-3.4989398,
similarity=46.198997]
5: 759317 Zimbabwe_national_cricket_team
[pem_pos=1, final_scores=0.914722,
final_local_scores=0.9020969, pem=-2.2776897,
similarity=43.396515]
6: 28357832 Singer_World_Series
[pem_pos=0, final_scores=0.6143275,
final_local_scores=0.68987185, pem=0.0, similarity=24.676865]

Sample: b'447_CRICKET'

CRICKET - [1 1 AUSTRALIA 1] WIN TOSS , OPT TO BAT AGAINST [2 2 ZIMBABWE 2] . [3 3 COLOMBO 3] 1996-08-26 [4 4 Australia 4] won the toss and decided to bat against [5 5 Zimbabwe 5] in the first day-night limited overs match of the [6 6 Singer World Series 6] tournament on Monday . Teams : [7 7 Australia 7] - [8 8 Mark Waugh 8] , [9 9 Michael Slater 9] , [10 10 Ricky Ponting 10] , [11 11 Steve Waugh 11] , [12 12 Stuart Law 12] , [13 13 Michael Bevan 13] , [14 14 Ian Healy 14] (captain) , [15 15 Brad Hogg 15] , [16 16 Paul Reiffel 16] , Damein Fleming , [17 17 Glenn McGrath 17] . [18 18 Zimbabwe 18] - [19 19 Alistair Campbell 19] (captain) , [20 20 Andy Flower 20] , [21 21 Grant Flower 21] , [22 22 Guy Whittall 22] , [23 23 Craig Evans 23] , [24 24 Eddo Brandes 24] , [25 25 Heath Streak 25] , [26 26 Paul Strang 26] , [27 27 Craig Wishart 27] , [28

```
28 Ali_Shah_28] , [29 29 Mark_Dekker_29] .
1: 374833 Australia_national_cricket_team
2: 759317 Zimbabwe_national_cricket_team
3: 56636 Colombo
4: 374833 Australia_national_cricket_team
5: 759317 Zimbabwe_national_cricket_team
6: 28357832 Singer_World_Series
7: 374833 Australia_national_cricket_team
8: 151606 Mark_Waugh
9: 896423 Michael_Slater
10: 462001 Ricky_Ponting
11: 156123 Steve_Waugh
12: 2047175 Stuart_Law
13: 284900 Michael_Bevan
14: 928267 Ian_Healy
15: 1384007 Brad_Hogg
16: 2383377 Paul_Reiffel
17: 423488 Glenn_McGrath
18: 759317 Zimbabwe_national_cricket_team
19: 2156908 Alistair_Campbell_(cricketer)
20: 733585 Andy_Flower
21: 897202 Grant_Flower
22: 2195319 Guy_Whittall
23: 2877538 Craig_Evans_(Zimbabwean_sportsman)
24: 3966081 Eddo_Brandes
25: 897195 Heath_Streak
26: 3751758 Paul_Strang
27: 2878441 Craig_Wishart
28: 4036464 Ali_Shah
29: 3022177 Mark_Dekker
1: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.65107745,
final_local_scores=0.6175506, pem=-3.4989398,
similarity=41.546432]
2: 759317 Zimbabwe_national_cricket_team
    [pem_pos=1, final_scores=0.66875,
final_local_scores=0.64699227, pem=-2.2776897,
similarity=35.783684]
3: 56636 Colombo
    [pem_pos=0, final_scores=0.507612,
final_local_scores=0.59006256, pem=-0.19393882,
similarity=22.753027]
4: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.7776168,
final_local_scores=0.7557757, pem=-3.4989398,
similarity=45.671345]
5: 759317 Zimbabwe_national_cricket_team
    [pem_pos=1, final_scores=0.91342384,
final_local_scores=0.91426116, pem=-2.2776897,
similarity=43.759518]
6: 28357832 Singer_World_Series
    [pem_pos=0, final_scores=0.6450944,
```

```
final_local_scores=0.71493995, pem=0.0, similarity=25.424948]
7: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.8131655,
final_local_scores=0.79460716, pem=-3.4989398,
similarity=46.830154]
8: 151606 Mark_Waugh
    [pem_pos=0, final_scores=0.5832384,
final_local_scores=0.5656404, pem=0.0, similarity=20.969551]
9: 896423 Michael_Slater
    [pem_pos=0, final_scores=0.53194463,
final_local_scores=0.55176556, pem=0.0, similarity=20.555498]
10: 462001 Ricky_Ponting
    [pem_pos=0, final_scores=0.50742275,
final_local_scores=0.47189933, pem=0.0, similarity=18.17213]
11: 156123 Steve_Waugh
    [pem_pos=0, final_scores=0.59295183,
final_local_scores=0.5730374, pem=0.0, similarity=21.190292]
12: 2047175 Stuart_Law
    [pem_pos=0, final_scores=0.8786132,
final_local_scores=0.91428727, pem=0.0, similarity=31.373867]
13: 284900 Michael_Bevan
    [pem_pos=0, final_scores=0.76734644,
final_local_scores=0.7861888, pem=0.0, similarity=27.551155]
14: 928267 Ian_Healy
    [pem_pos=0, final_scores=0.8488224,
final_local_scores=0.8652908, pem=0.0, similarity=29.911716]
15: 1384007 Brad_Hogg
    [pem_pos=0, final_scores=0.9627226,
final_local_scores=0.99149245, pem=0.0, similarity=33.677826]
16: 2383377 Paul_Reiffel
    [pem_pos=0, final_scores=0.6640267,
final_local_scores=0.695619, pem=0.0, similarity=24.848372]
17: 423488 Glenn_McGrath
    [pem_pos=0, final_scores=0.8125332,
final_local_scores=0.8072948, pem=0.0, similarity=28.181]
18: 759317 Zimbabwe_national_cricket_team
    [pem_pos=1, final_scores=0.59773284,
final_local_scores=0.5694167, pem=-2.2776897,
similarity=33.468674]
19: 2156908 Alistair_Campbell_(cricketer)
    [pem_pos=0, final_scores=0.6837358,
final_local_scores=0.7135982, pem=-0.4234836,
similarity=27.687876]
20: 733585 Andy_Flower
    [pem_pos=0, final_scores=0.89866203,
final_local_scores=0.9122848, pem=0.0, similarity=31.314112]
21: 897202 Grant_Flower
    [pem_pos=0, final_scores=0.9236682,
final_local_scores=0.95749253, pem=0.0, similarity=32.6632]
22: 2195319 Guy_Whittall
    [pem_pos=0, final_scores=0.553009,
final_local_scores=0.5836713, pem=0.0, similarity=21.507627]
```

```
23: 2877538 Craig_Evans_(Zimbabwean_sportsman)
    [pem_pos=3, final_scores=0.38799453,
final_local_scores=0.4709785, pem=-1.89712,
similarity=28.461485]
24: 3966081 Eddo_Brandes
    [pem_pos=0, final_scores=0.7953991,
final_local_scores=0.88977116, pem=0.0, similarity=30.642258]
25: 897195 Heath_Streak
    [pem_pos=0, final_scores=0.9747358,
final_local_scores=1.0224788, pem=0.0, similarity=34.60252]
26: 3751758 Paul_Strang
    [pem_pos=0, final_scores=0.8125579,
final_local_scores=0.87838584, pem=0.0, similarity=30.3025]
27: 2878441 Craig_Wishart
    [pem_pos=0, final_scores=0.71239954,
final_local_scores=0.78778833, pem=0.0, similarity=27.598888]
28: 4036464 Ali_Shah
    [pem_pos=0, final_scores=0.81216604,
final_local_scores=0.93771845, pem=0.0, similarity=32.0731]
29: 3022177 Mark_Dekker
    [pem_pos=0, final_scores=0.44478363,
final_local_scores=0.5519386, pem=0.0, similarity=20.560661]
```

Sample: b'448_PRESS'

PRESS DIGEST - [1 1 MOZAMBIQUE 1] - AUGUST 26 . [2 2 MAPUTO 2] 1996-08-26 This is the leading story in the [3 3 Mozambican 3] press on Monday . [1 Reuters 1] has not verified this story and does not vouch for its accuracy . NOTICIAS - At least 20 people were killed when the two trucks in which they were travelling collided at Nhamavila about 160 km north of [4 4 Maputo 4] on Saturday , the [5 5 Maputo 5] daily Noticias said .

```
1: 19301 Mozambique
2: 57044 Maputo
3: 19301 Mozambique
4: 57044 Maputo
5: 57044 Maputo
1: 19301 Mozambique
    [pem_pos=0, final_scores=0.47271383,
final_local_scores=0.4468248, pem=-0.36112452,
similarity=19.387707]
2: 57044 Maputo
    [pem_pos=0, final_scores=0.46931922,
final_local_scores=0.4144373, pem=0.0, similarity=16.457348]
3: 19301 Mozambique
    [pem_pos=0, final_scores=0.3701262,
final_local_scores=0.33476332, pem=0.0, similarity=14.079718]
4: 57044 Maputo
    [pem_pos=0, final_scores=0.47507542,
final_local_scores=0.4207251, pem=0.0, similarity=16.644989]
5: 57044 Maputo
    [pem_pos=0, final_scores=0.3834893,
```

```
final_local_scores=0.32068112, pem=0.0, similarity=13.659477]
1: 18998750 Reuters
    [pem_pos=0, final_scores=0.21998586,
final_local_scores=0.3002629, pem=0.0, similarity=13.050157]
```

Sample: b'449_Lamonts'

Lamonts Apparel files reorganization plan . [CORRECTED
18:00 [1 1 GMT 1]] [2 2 KIRKLAND 2] , [3 3 Wash 3] .
1996-08-26 Lamonts Apparel Inc , an operator of 42 family apparel stores in six northwestern states , said it has filed a reorganization plan in bankruptcy court in [4 4 Seattle 4] . (Corrects to make clear a reorganization plan has been filed) . The reorganization plan calls for all secured claims to be paid in full . Unsecured claims , including those of company bondholders , will be satisfied by issuing new common stock and warrants . Unsecured claims are estimated at about \$ 90 million . [5 5 Lamonts 5] said it plans to issue 9 million shares of new common stock . Of that amount , 4.05 million and 5.67 million shares will be allocated to the company 's trade creditors . Between 4.75 and 3.13 million shares will be allocated to bondholders and other unsecured non-trade creditors and 200,000 will be allocated to existing shareholders in exchange for all existing stock of the company . Bondholders and other unsecured non-trade creditors will receive warrants four about 2.2 million shares when the company 's market capitalization reaches \$ 20 million . Bondholders , other unsecured non-trade creditors , and existing shareholders also will receive warrants entitling them to roughly 800,000 shares when the company 's market capitalization reaches \$ 25 million . Management will receive options to purchase 10 percent of the company 's outstanding common stock with protection against dilution at an option exercise price of \$ 1 million .

```
1: 12701 Greenwich_Mean_Time
2: 138007 Kirkland,_Washington
3: 13015878 Washington_(state)
4: 11388236 Seattle
5: 6036380 Lamonts
1: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.5596462,
final_local_scores=0.71071786, pem=0.0, similarity=25.298952]
2: 138007 Kirkland,_Washington
    [pem_pos=0, final_scores=0.53688234,
final_local_scores=0.5586453, pem=-0.7361297,
similarity=24.76399]
3: 13015878 Washington_(state)
    [pem_pos=6, final_scores=0.63061506,
final_local_scores=0.6621582, pem=-2.76717,
similarity=38.898136]
4: 11388236 Seattle
    [pem_pos=0, final_scores=0.43312395,
```

```
final_local_scores=0.4096133, pem=-0.1054736,
similarity=16.886974]
5: 6036380 Lamonts
    [pem_pos=0, final_scores=0.3641643,
final_local_scores=0.47046685, pem=0.0, similarity=18.129383]

Sample: b'450_Passengers'
Passengers rescued from blazing ferry off [1 1 France 1] . [2
2 LONDON 2] 1996-08-26 More than 100 people were safely
evacuated on Monday from a ferry that caught fire soon after
leaving [3 3 Guernsey 3] in [4 4 Britain 4] 's [5 5 Channel Is
lands 5] , police said . Police said the 111 passengers and
six crew on board the ferry Trident Seven , owned by [6 6 Fra
nce 6] 's Emeraud line , were rescued by a variety of private
and commercial boats after fire broke out in the engine room
soon after it left port . An 88-year-old woman was taken to
hospital with leg injuries , according to a spokesman for [7 7
Guernsey 7] police . The ferry , which was towed into port ,
had been bound for [8 8 Jersey 8] , another in a cluster of
small British-ruled islands off north-west [9 9 France 9] .
1: 5843419 France
2: 17867 London
3: 12166 Guernsey
4: 31717 United_Kingdom
5: 5643 Channel_Islands
6: 5843419 France
7: 12166 Guernsey
8: 15693 Jersey
9: 5843419 France
1: 5843419 France
    [pem_pos=0, final_scores=0.61360127,
final_local_scores=0.5861788, pem=-0.16240834,
similarity=22.465658]
2: 17867 London
    [pem_pos=0, final_scores=0.7115223,
final_local_scores=0.7223218, pem=-0.09086737,
similarity=26.139387]
3: 12166 Guernsey
    [pem_pos=0, final_scores=0.42629892,
final_local_scores=0.3806507, pem=-0.31114414,
similarity=17.141138]
4: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.4925602,
final_local_scores=0.46778768, pem=-0.61114657,
similarity=21.37294]
5: 5643 Channel_Islands
    [pem_pos=0, final_scores=0.34094393,
final_local_scores=0.295907, pem=-0.15505362,
similarity=13.763373]
6: 5843419 France
    [pem_pos=0, final_scores=0.6736919,
final_local_scores=0.65181875, pem=-0.16240834,
```

```
similarity=24.424486]
7: 12166 Guernsey
    [pem_pos=0, final_scores=0.42439896,
final_local_scores=0.37857527, pem=-0.31114414,
similarity=17.079205]
8: 15693 Jersey
    [pem_pos=0, final_scores=0.49271196,
final_local_scores=0.4546618, pem=-0.24350123,
similarity=18.981926]
9: 5843419 France
    [pem_pos=0, final_scores=0.67364067,
final_local_scores=0.6517627, pem=-0.16240834,
similarity=24.422813]
```

Sample: b'451_OSCE'

OSCE delays decision on refugee voting . [1 1 Kurt Schork 1] [2 2 SARAJEVO 2] 1996-08-26 Bosnia 's election organisers will decide on Tuesday whether or not to postpone municipal elections scheduled as part of nationwide balloting , an OSCE spokeswoman said on Monday . Officials from the Organisation for Security and Cooperation in Europe (OSCE) are considering the postponement following allegations of serious irregularities in the registration of [3 3 Serb 3] refugees . International observers say the alleged irregularities could affect the outcome of voting for municipal assemblies . " Tomorrow ... the Provisional Election Commission will consider the possible postponement of municipal elections only ... the other elections will be held on September 14 , " OSCE spokeswoman Agota Kuperman told reporters in [4 4 Sarajevo 4] . " I think that it would be very difficult to select which municipal elections would have to be cancelled . I think probably if the decision (to cancell) were to be taken it would probably be all municipal elections . . . " Kuperman added that options other than postponement were also on the table , but she refused to specify what they were . The 1 Dayton peace agreement gave the OSCE a mandate to organise Bosnian elections . The Provisional Election Commission is OSCE 's top rule-making body for the poll . More than 600,000 refugees have registered to vote in 55 countries around the world , representing about 20 per cent of Bosnia 's total electorate . They are due to begin voting on Wednesday , August 28 , just one day after the PEC is supposed to make its decision . Balloting inside Bosnia is scheduled for September 14 , when citizens are slated to elect municipal and cantonal assemblies , separate Moslem-Croat and [5 5 Serb 5] parliaments , a national [6 6 House of Representatives 6] and a three-man Presidency . A [7 7 Sarajevo 7] daily newspaper , Dnevi Avaz , which is close to Bosnia 's [8 8 Moslem 8] nationalist [9 9 SDA 9] party , on Monday said the OSCE would postpone the municipal-level elections until the Spring of 1997 because of the refugee registration problems . [10 10 SDA 10] has a representative on the

Provisional Election Commission . " I do n't know what the source of the Dnevi Avaz report is , but it is consistent with what I have heard from western diplomats and from inside the OSCE , " said an OSCE staff member in [11 11 Sarajevo 11] who asked not to be named . " The word is that Frowick has decided to postpone municipal elections but that he will wait for one more session of the PEC on Tuesday to take everyone 's temperature on the issue . " Ambassador Robert Frowick , an [1 2 American 1] , heads the OSCE mission in Bosnia . OSCE and independent monitors allege that [12 12 Serb 12] authorities have systematically discouraged refugees from registering to cast a ballot in the places they lived before the war . Instead , the refugees were said to have been directed by their authorities to vote from strategic towns which had [13 13 Moslem 13] majorities before the 43-month [14 14 Bosnian 14] war , but which are now underpopulated as a result of " ethnic cleansing " . Diplomats explain the purpose of this electoral engineering is to secure [15 15 Serb 15] control over pivotal towns inside the 49 per cent of Bosnia known as the [16 16 Serb 16] republic , consolidating through the ballot box what was initially taken in war . The top [17 17 U.N. 17] refugee official in the [18 18 Balkans 18] highlighted the voter registration problem on Monday . " Results of the registration for September elections herald a dismal future for multi-ethnicity in Bosnia-Herzegovina , " warned [19 19 Soren Jessen-Petersen 19] , Special Envoy for the [20 20 U.N. 20] High Commissioner for Refugees in Former Yugoslavia . " In the run-up to elections , nationalistic political leaders are playing the ethnic / sectarian card , drumming up support within their constituencies by playing on bitter memories or fear . "

1: 4378291 Kurt_Schork
2: 26786 Sarajevo
3: 29265 Serbia
4: 26786 Sarajevo
5: 29265 Serbia
6: 19468532 House_of_Representatives
7: 26786 Sarajevo
8: 6037917 Islam
9: 349662 Party_of_Democratic_Action
10: 349662 Party_of_Democratic_Action
11: 26786 Sarajevo
12: 29265 Serbia
13: 6037917 Islam
14: 3114713 Bosnia_(region)
15: 29265 Serbia
16: 29265 Serbia
17: 31769 United_Nations
18: 4829 Balkans
19: 3795700 Søren_Jessen-Petersen
20: 31769 United_Nations
1: 168269 Dayton_Agreement

```
[winner: 8253 Dayton,_Ohio, pem_pos=0,
final_scores=-0.039213657, final_local_scores=-0.024786696,
pem=-0.9297534, similarity=8.406172]
    RECALL MISS: 168269 Dayton_Agreement]
2: 19792942 Americans
    [winner: 3434750 United_States, pem_pos=0,
final_scores=0.16228037, final_local_scores=0.14850548,
pem=-0.12994486, similarity=9.228073]
    gt: 19792942 Americans, pem_pos=1,
final_scores=0.006571695, final_local_scores=0.05131322,
pem=-3.7045176, similarity=25.766745 ]
1: 4378291 Kurt_Schork
    [pem_pos=0, final_scores=0.5473789,
final_local_scores=0.6876803, pem=0.0, similarity=24.611465]
2: 26786 Sarajevo
    [pem_pos=0, final_scores=0.4103362,
final_local_scores=0.36915672, pem=-0.0017846047,
similarity=15.11579]
3: 29265 Serbia
    [pem_pos=1, final_scores=0.6306558,
final_local_scores=0.60789526, pem=-2.1726613,
similarity=34.045788]
4: 26786 Sarajevo
    [pem_pos=0, final_scores=0.4091279,
final_local_scores=0.3678369, pem=-0.0017846047,
similarity=15.076405]
5: 29265 Serbia
    [pem_pos=1, final_scores=0.518431,
final_local_scores=0.48530662, pem=-2.1726613,
similarity=30.387499]
6: 19468532 House_of_Representatives
    [pem_pos=5, final_scores=0.45503014,
final_local_scores=0.58057123, pem=-3.0198689,
similarity=37.83763]
7: 26786 Sarajevo
    [pem_pos=0, final_scores=0.4214859,
final_local_scores=0.3813361, pem=-0.0017846047,
similarity=15.479248]
8: 6037917 Islam
    [pem_pos=1, final_scores=0.40316457,
final_local_scores=0.4475723, pem=-1.2150909,
similarity=24.054016]
9: 349662 Party_of_Democratic_Action
    [pem_pos=0, final_scores=0.42220443,
final_local_scores=0.49490577, pem=-0.3944092,
similarity=21.003548]
10: 349662 Party_of_Democratic_Action
    [pem_pos=0, final_scores=0.41660708,
final_local_scores=0.48879153, pem=-0.3944092,
similarity=20.821085]
11: 26786 Sarajevo
    [pem_pos=0, final_scores=0.4340971,
```

```
final_local_scores=0.39511192, pem=-0.0017846047,
similarity=15.890346]
12: 29265 Serbia
    [pem_pos=1, final_scores=0.5826469,
final_local_scores=0.5554527, pem=-2.1726613,
similarity=32.4808]
13: 6037917 Islam
    [pem_pos=1, final_scores=0.3984403,
final_local_scores=0.44241178, pem=-1.2150909,
similarity=23.900015]
14: 3114713 Bosnia_(region)
    [pem_pos=5, final_scores=0.16241817,
final_local_scores=0.15505034, pem=-2.2460418,
similarity=20.931047]
15: 29265 Serbia
    [pem_pos=1, final_scores=0.5456138,
final_local_scores=0.51499975, pem=-2.1726613,
similarity=31.273603]
16: 29265 Serbia
    [pem_pos=1, final_scores=0.614325,
final_local_scores=0.5900563, pem=-2.1726613,
similarity=33.51344]
17: 31769 United_Nations
    [pem_pos=0, final_scores=0.4499303,
final_local_scores=0.4672746, pem=0.0, similarity=18.034119]
18: 4829 Balkans
    [pem_pos=0, final_scores=0.4989378,
final_local_scores=0.51885194, pem=0.0, similarity=19.57329]
19: 3795700 Søren_Jessen-Petersen
    [pem_pos=0, final_scores=0.40935022,
final_local_scores=0.5332088, pem=0.0, similarity=20.001726]
20: 31769 United_Nations
    [pem_pos=0, final_scores=0.52361757,
final_local_scores=0.54776675, pem=0.0, similarity=20.436165]
1: 3434750 United_States
    [pem_pos=0, final_scores=0.16228037,
final_local_scores=0.14850548, pem=-0.12994486,
similarity=9.228073]
```

Sample: b'452_After'

After truce , [1 1 Lebed 1] faces tougher [2 2 Chechen 2] problem . Alastair Macdonald [3 3 MOSCOW 3] 1996-08-26 [4 4 Alexander Lebed 4] may finally get to discuss his [5 5 Chechen 5] peace proposals with [6 6 Boris Yeltsin 6] on Tuesday after a lost weekend in the region when he was forced to abandon plans to sign a new political treaty with the separatist rebels . It could be a sobering experience for the [7 7 Kremlin in 7] security chief who , just two months after being appointed by the president and two heady weeks after taking charge of the [8 8 Chechen 8] crisis , had promised to wrap up the 20-month war by the weekend . Despite a mood of compromise in the region after some of the worst fighting of

the war , [9 9 Lebed 9] may be just finding out that concluding a long-term settlement , somewhere between rebel demands for independence and [10 10 Moscow 10] 's insistence that [11 11 Chechnya 11] remain part of [12 12 Russia 12] , will be no easy matter . A chain-smoking former paratroop general with a sharp line in deadpan putdowns and a soldier 's knack for making life sound simple , [13 13 Lebed 13] managed to arrange an ambitious ceasefire in the region last week , days after the [14 14 Russian 14] army threatened to bomb its way back into the rebel-held [15 15 Chechen 15] capital [16 16 Grozny 16] . He returned at the weekend , pledging to conclude a political settlement that would make this truce work where all others have failed . But then he apparently thought better of it . Saying he needed to tidy up legal loose ends on the deal -- and also cover his back against unnamed pro-war schemers in [17 17 Moscow 17] -- he flew back to the capital empty-handed on Sunday . He met [18 1 8 Russian 18] Prime Minister [19 19 Viktor Chernomyrdin 19] on Monday and , according to the press service of [20 20 Lebed 20] 's [21 21 Security Council 21] , could meet [22 22 Yeltsin 22] on Tuesday . It was not clear if the start of [23 23 Yelt sin 23] 's holiday , announced later , would affect plans to talk . 1 Moscow 's as yet undisclosed proposals on [24 24 Chechnya 24] 's political future have , meanwhile , been sent back to do the rounds of various government departments . It is not known what caused the delay in the peace plan . Speculation mounted last week that [25 25 Lebed 25] was operating out on a limb in [26 26 Chechnya 26] as [27 27 Yelt sin 27] , hardly seen since his reelection last month , kept out of sight and then gave an interview criticising his envoy just as he clinched a truce . But [28 28 Chernomyrdin 28] took pains at the weekend to insist [29 29 Lebed 29] was playing for a united team and that the proposals he took to rebel chief-of-staff [30 30 Aslan Maskhadov 30] had been agreed by [31 31 Yeltsin 31] . If that were the case , it was unclear why [32 32 Lebed 32] suddenly found it necessary to have the deal verified in [33 33 Moscow 33] . [34 34 Lebed 34] himself said he was concerned that powerful interests in [35 3 5 Moscow 35] , who have profited from the war and want it to go on , would seize on any inadequacy in the deal to remove him . [36 36 Maskhadov 36] warned that [37 37 Lebed 37] risked the same fate as a former [38 38 Russian 38] army commander in the region who sought a compromise and was blown up last year by -- so the [39 39 Chechens 39] say -- [40 40 Russian 40] forces . [41 41 Maskhadov 41] may also have made new demands on [42 42 Lebed 42] , forcing him to check back with the [43 43 Kremlin 43] on what he could offer . A rebel spokesman said the two military men had come close to agreement on a face-saving formula acceptable to both sides and involving a referendum on independence at some future date . [44 44 Chernomyrdin 44] said voters should decide [45 4 5 Chechnya 45] 's future but stressed there was no question

of the government letting the region quit the [46 46 Russian Federation 46] -- something [47 47 Moscow 47] fears could encourage separatist tendencies in other ethnic regions , particularly in the strategic [48 48 North Caucasus 48] . [49 49 Lebed 49] may be be finding that closing that political gap between the two sides is more difficult than just ending a war .

1: 431579 Alexander_Lebed
2: 6095 Chechnya
3: 19004 Moscow
4: 431579 Alexander_Lebed
5: 6095 Chechnya
6: 23278520 Boris_Yeltsin
7: 381075 Moscow_Kremlin
8: 6095 Chechnya
9: 431579 Alexander_Lebed
10: 19004 Moscow
11: 6095 Chechnya
12: 25391 Russia
13: 431579 Alexander_Lebed
14: 25391 Russia
15: 6095 Chechnya
16: 140790 Grozny
17: 19004 Moscow
18: 25391 Russia
19: 482420 Viktor_Chernomyrdin
20: 431579 Alexander_Lebed
21: 7516389 Security_Council_of_Russia
22: 23278520 Boris_Yeltsin
23: 23278520 Boris_Yeltsin
24: 6095 Chechnya
25: 431579 Alexander_Lebed
26: 6095 Chechnya
27: 23278520 Boris_Yeltsin
28: 482420 Viktor_Chernomyrdin
29: 431579 Alexander_Lebed
30: 340963 Aslan_Maskhadov
31: 23278520 Boris_Yeltsin
32: 431579 Alexander_Lebed
33: 19004 Moscow
34: 431579 Alexander_Lebed
35: 19004 Moscow
36: 340963 Aslan_Maskhadov
37: 431579 Alexander_Lebed
38: 25391 Russia
39: 6095 Chechnya
40: 25391 Russia
41: 340963 Aslan_Maskhadov
42: 431579 Alexander_Lebed
43: 9787373 Government_of_Russia
44: 482420 Viktor_Chernomyrdin
45: 6095 Chechnya

46: 25391 Russia
47: 19004 Moscow
48: 482121 North_Caucasus
49: 431579 Alexander_Lebed
1: 9787373 Government_of_Russia
 [winner: 19004 Moscow, pem_pos=0,
final_scores=-0.19331814, final_local_scores=-0.25708655,
pem=-0.106582455, similarity=-3.0026507]
 RECALL MISS: 9787373 Government_of_Russia]
1: 431579 Alexander_Lebed
 [pem_pos=0, final_scores=0.65260047,
final_local_scores=0.6232062, pem=0.0, similarity=22.68743]
2: 6095 Chechnya
 [pem_pos=2, final_scores=0.5112832,
final_local_scores=0.49117124, pem=-0.8617233,
similarity=23.433428]
3: 19004 Moscow
 [pem_pos=0, final_scores=0.6647891,
final_local_scores=0.6831441, pem=-0.106582455,
similarity=25.055708]
4: 431579 Alexander_Lebed
 [pem_pos=0, final_scores=0.62093675,
final_local_scores=0.58721024, pem=0.0, similarity=21.613237]
5: 6095 Chechnya
 [pem_pos=2, final_scores=0.49644405,
final_local_scores=0.4749617, pem=-0.8617233,
similarity=22.949703]
6: 23278520 Boris_Yeltsin
 [pem_pos=0, final_scores=0.6549055,
final_local_scores=0.6387307, pem=0.0, similarity=23.150711]
7: 381075 Moscow_Kremlin
 [pem_pos=1, final_scores=0.30857414,
final_local_scores=0.32555425, pem=-0.76193243,
similarity=17.948408]
8: 6095 Chechnya
 [pem_pos=2, final_scores=0.5050934,
final_local_scores=0.48440987, pem=-0.8617233,
similarity=23.231655]
9: 431579 Alexander_Lebed
 [pem_pos=0, final_scores=0.7234805,
final_local_scores=0.70063186, pem=0.0, similarity=24.997965]
10: 19004 Moscow
 [pem_pos=0, final_scores=0.5208147,
final_local_scores=0.52587396, pem=-0.106582455,
similarity=20.362453]
11: 6095 Chechnya
 [pem_pos=0, final_scores=0.57436925,
final_local_scores=0.5600832, pem=0.0, similarity=20.803713]
12: 25391 Russia
 [pem_pos=0, final_scores=0.78381234,
final_local_scores=0.8289235, pem=-0.20299491,
similarity=29.93036]

13: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.6467516,
final_local_scores=0.6168172, pem=0.0, similarity=22.496767]
14: 25391 Russia
[pem_pos=1, final_scores=0.8576588,
final_local_scores=0.90958947, pem=-1.0515022,
similarity=36.951912]
15: 6095 Chechnya
[pem_pos=2, final_scores=0.49227685,
final_local_scores=0.47040975, pem=-0.8617233,
similarity=22.813864]
16: 140790 Grozny
[pem_pos=0, final_scores=0.520608,
final_local_scores=0.5198565, pem=0.0, similarity=19.603268]
17: 19004 Moscow
[pem_pos=0, final_scores=0.56897295,
final_local_scores=0.5784795, pem=-0.106582455,
similarity=21.932308]
18: 25391 Russia
[pem_pos=1, final_scores=0.8869545,
final_local_scores=0.9415906, pem=-1.0515022,
similarity=37.906887]
19: 482420 Viktor_Chernomyrdin
[pem_pos=0, final_scores=0.80653745,
final_local_scores=0.84611654, pem=0.0, similarity=29.339518]
20: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.6770964,
final_local_scores=0.64996433, pem=0.0, similarity=23.485945]
21: 7516389 Security_Council_of_Russia
[pem_pos=1, final_scores=0.23015068,
final_local_scores=0.23529816, pem=-1.0463139,
similarity=16.801497]
22: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.57925624,
final_local_scores=0.55609536, pem=0.0, similarity=20.684708]
23: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.59586626,
final_local_scores=0.5742393, pem=0.0, similarity=21.22616]
24: 6095 Chechnya
[pem_pos=0, final_scores=0.5917267,
final_local_scores=0.57904357, pem=0.0, similarity=21.369528]
25: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.6809694,
final_local_scores=0.65419495, pem=0.0, similarity=23.612196]
26: 6095 Chechnya
[pem_pos=0, final_scores=0.54861087,
final_local_scores=0.53194606, pem=0.0, similarity=19.964045]
27: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.58080107,
final_local_scores=0.5577828, pem=0.0, similarity=20.735065]
28: 482420 Viktor_Chernomyrdin
[pem_pos=0, final_scores=0.60259557,

```
final_local_scores=0.62334096, pem=0.0, similarity=22.69145]
29: 431579 Alexander_Lebed
    [pem_pos=0, final_scores=0.70188445,
final_local_scores=0.6770415, pem=0.0, similarity=24.293982]
30: 340963 Aslan_Maskhadov
    [pem_pos=0, final_scores=0.708402,
final_local_scores=0.71908325, pem=0.0, similarity=25.548592]
31: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.611183,
final_local_scores=0.5909705, pem=0.0, similarity=21.725452]
32: 431579 Alexander_Lebed
    [pem_pos=0, final_scores=0.72031134,
final_local_scores=0.69717, pem=0.0, similarity=24.894657]
33: 19004 Moscow
    [pem_pos=0, final_scores=0.53974134,
final_local_scores=0.5465484, pem=-0.106582455,
similarity=20.97942]
34: 431579 Alexander_Lebed
    [pem_pos=0, final_scores=0.7134685,
final_local_scores=0.68969524, pem=0.0, similarity=24.671595]
35: 19004 Moscow
    [pem_pos=0, final_scores=0.49085635,
final_local_scores=0.49314898, pem=-0.106582455,
similarity=19.385874]
36: 340963 Aslan_Maskhadov
    [pem_pos=0, final_scores=0.5155475,
final_local_scores=0.50841904, pem=0.0, similarity=19.261951]
37: 431579 Alexander_Lebed
    [pem_pos=0, final_scores=0.6760891,
final_local_scores=0.6488639, pem=0.0, similarity=23.453106]
38: 25391 Russia
    [pem_pos=1, final_scores=0.8394136,
final_local_scores=0.88965935, pem=-1.0515022,
similarity=36.357155]
39: 6095 Chechnya
    [pem_pos=2, final_scores=0.50695264,
final_local_scores=0.48644078, pem=-0.96860826,
similarity=23.873518]
40: 25391 Russia
    [pem_pos=1, final_scores=0.86481076,
final_local_scores=0.91740197, pem=-1.0515022,
similarity=37.18505]
41: 340963 Aslan_Maskhadov
    [pem_pos=0, final_scores=0.47854906,
final_local_scores=0.46800387, pem=0.0, similarity=18.055882]
42: 431579 Alexander_Lebed
    [pem_pos=0, final_scores=0.7087458,
final_local_scores=0.68453646, pem=0.0, similarity=24.517647]
43: 9787373 Government_of_Russia
    [pem_pos=3, final_scores=0.48978603,
final_local_scores=0.48901314, pem=-2.0052693,
similarity=29.587807]
```

```
44: 482420 Viktor_Chernomyrdin
    [pem_pos=0, final_scores=0.6637727,
final_local_scores=0.69016767, pem=0.0, similarity=24.685692]
45: 6095 Chechnya
    [pem_pos=0, final_scores=0.532026,
final_local_scores=0.5138296, pem=0.0, similarity=19.423412]
46: 25391 Russia
    [pem_pos=0, final_scores=0.5439723,
final_local_scores=0.56693465, pem=0.0, similarity=21.008173]
47: 19004 Moscow
    [pem_pos=0, final_scores=0.45188528,
final_local_scores=0.450579, pem=-0.106582455,
similarity=18.115501]
48: 482121 North_Caucasus
    [pem_pos=0, final_scores=0.3542462,
final_local_scores=0.3490292, pem=0.0, similarity=14.505441]
49: 431579 Alexander_Lebed
    [pem_pos=0, final_scores=0.6394084,
final_local_scores=0.6087959, pem=0.0, similarity=22.257397]
```

Sample: b'453_Yeltsin'

[1 1 Yeltsin 1] on holiday from Monday - [2 2 Interfax 2] . [3 3 MOSCOW 3] 1996-08-26 [4 4 Russian 4] President [5 5 Boris Yeltsin 5] began a new summer holiday on Monday but would remain in control of affairs of state , [6 6 Interfax 6] news agency said . " The head of state 's holiday has only just begun , " the agency quoted [7 7 Sergei Yastrzhembsky 7] as saying , adding that the president was currently in a [8 8 Kremlin 8] residence near [9 9 Moscow 9] . [10 10 Interfax 10] said Yastrezhembsky did not exclude that [11 11 Yeltsin 11] could spend some time in other places . He would continue working on various documents and might meet " one state official or another " . " One must give B. Yeltsin a chance to rest and recover his health after the elections , " [12 12 Interfax 12] quoted Yastrezhembsky as saying . " ([13 13 Yeltsin 13]) controls internal and international policies , daily receives a big packet of documents from [14 14 Moscow 14] , which demand his intervention ... Many of those documents return to the president 's administration the same day , " he said . [15 15 Yeltsin 15] went on a two-day trip outside [16 16 Moscow 16] last week to check out a holiday home .

```
1: 23278520 Boris_Yeltsin
2: 4921404 Interfax
3: 19004 Moscow
4: 25391 Russia
5: 23278520 Boris_Yeltsin
6: 4921404 Interfax
7: 13411146 Sergey_Yastrzhembsky
8: 381075 Moscow_Kremlin
9: 19004 Moscow
10: 4921404 Interfax
11: 23278520 Boris_Yeltsin
```

```
12: 4921404 Interfax
13: 23278520 Boris_Yeltsin
14: 19004 Moscow
15: 23278520 Boris_Yeltsin
16: 19004 Moscow
1: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.46043962,
final_local_scores=0.41573876, pem=0.0, similarity=16.496187]
2: 4921404 Interfax
    [pem_pos=0, final_scores=0.46902364,
final_local_scores=0.46327102, pem=0.0, similarity=17.914644]
3: 19004 Moscow
    [pem_pos=0, final_scores=0.7206708,
final_local_scores=0.7202804, pem=-0.106582455,
similarity=26.163929]
4: 25391 Russia
    [pem_pos=1, final_scores=0.7738263,
final_local_scores=0.81952596, pem=-1.0515022,
similarity=34.264233]
5: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.6844457,
final_local_scores=0.6604313, pem=0.0, similarity=23.798302]
6: 4921404 Interfax
    [pem_pos=0, final_scores=0.7338681,
final_local_scores=0.75257343, pem=0.0, similarity=26.548006]
7: 13411146 Sergey_Yastrzhembsky
    [pem_pos=0, final_scores=0.5372313,
final_local_scores=0.6084005, pem=0.0, similarity=22.245598]
8: 381075 Moscow_Kremlin
    [pem_pos=1, final_scores=0.40733486,
final_local_scores=0.4133116, pem=-0.76193243,
similarity=20.567263]
9: 19004 Moscow
    [pem_pos=0, final_scores=0.60007125,
final_local_scores=0.58854365, pem=-0.106582455,
similarity=22.232641]
10: 4921404 Interfax
    [pem_pos=0, final_scores=0.60581124,
final_local_scores=0.61269075, pem=0.0, similarity=22.373627]
11: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.61819917,
final_local_scores=0.58806705, pem=0.0, similarity=21.638807]
12: 4921404 Interfax
    [pem_pos=0, final_scores=0.6499782,
final_local_scores=0.6609365, pem=0.0, similarity=23.813375]
13: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.66220623,
final_local_scores=0.6361381, pem=0.0, similarity=23.073341]
14: 19004 Moscow
    [pem_pos=0, final_scores=0.39695746,
final_local_scores=0.36667275, pem=-0.106582455,
similarity=15.611572]
```

```
15: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.62880224,
final_local_scores=0.59964937, pem=0.0, similarity=21.984446]
16: 19004 Moscow
    [pem_pos=0, final_scores=0.5830088,
final_local_scores=0.5699056, pem=-0.106582455,
similarity=21.676445]
```

Sample: b'454_Yugo'

Yugo Zastava workers ' protest enters 2nd week . [1 1 BELGRA
DE 1] 1996-08-26 Thousands of workers of [2 2 Serbia 2] 's [3
3 Zastava 3] arms factory entered the second week of protests
on Monday over unpaid wages and the lack of a programme to
revive the plant 's production . " We are stubborn , have
strength and time to persist until our demands are met , "
said the factory 's trade union secretary Dragutin
Stanojlovic . " We are united and we are waiting for the
government to decide what to do with us . " Trade unions
demanded payment of June and July wages and last year 's
holiday pay , and called on government to develop a revival
programme for the plant . The former [4 4 Yugoslav 4]
national army consumed 90 percent of [5 5 Zastava 5] 's pre-
war output , but like the rest of [6 6 Yugoslavia 6] 's
economy , the new army of [7 7 Serbia 7] and [8 8 Montenegro 8]
] is crippled by lack of funds . In [9 9 Kragujevac 9] where
the plant is based , 9,000 to 10,000 people gathered in the
central square to express their bitterness at what
Stanojlovic called government indifference . The ruling [10 10
Socialist Party 10] last week accused [11 11 Serbia 11] 's
opposition of stirring up social unrest and using workers '
already difficult social position for their own interests . "
Late wages , no work and (the lack of a) production
programme are the main reasons for their protests , " a
senior member of the local [12 12 Socialist Party 12] branch
told [13 13 Reuters 13] . " But there is a significant
influence of opposition parties that are collecting points
ahead of coming elections . " Federal [14 14 Yugoslav 14]
elections are due on November 3 . The party official said the
union 's figure on the number of protesters was exaggerated .
" There were about 1,400 to 1,600 people at the protest today
-- much less than last week . " But there were more observers
and passers by , " he said . -- Gordana Kukic , Belgrade
Newsroom +381 11 222 4254

```
1: 55904 Belgrade
2: 29265 Serbia
3: 8370927 Zastava_Arms
4: 34244 Yugoslavia
5: 8370927 Zastava_Arms
6: 34244 Yugoslavia
7: 29265 Serbia
8: 20760 Montenegro
9: 147111 Kragujevac
```

```
10: 417429 Socialist_Party_of_Serbia
11: 29265 Serbia
12: 417429 Socialist_Party_of_Serbia
13: 18998750 Reuters
14: 34244 Yugoslavia
1: 55904 Belgrade
    [pem_pos=0, final_scores=0.51579696,
final_local_scores=0.47534782, pem=-0.05193398,
similarity=18.557465]
2: 29265 Serbia
    [pem_pos=0, final_scores=0.6450407,
final_local_scores=0.61421984, pem=-0.34077588,
similarity=24.272451]
3: 8370927 Zastava_Arms
    [pem_pos=1, final_scores=0.48350704,
final_local_scores=0.5790147, pem=-0.64282215,
similarity=24.864433]
4: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.37990505,
final_local_scores=0.31567097, pem=-0.7596219,
similarity=17.640905]
5: 8370927 Zastava_Arms
    [pem_pos=1, final_scores=0.5076149,
final_local_scores=0.6053489, pem=-0.64282215,
similarity=25.6503]
6: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.43900806,
final_local_scores=0.38023204, pem=-0.6536551,
similarity=18.991274]
7: 29265 Serbia
    [pem_pos=0, final_scores=0.573335,
final_local_scores=0.53589237, pem=-0.34077588,
similarity=21.935003]
8: 20760 Montenegro
    [pem_pos=0, final_scores=0.5549142,
final_local_scores=0.5250213, pem=-0.44913077,
similarity=22.199839]
9: 147111 Kragujevac
    [pem_pos=0, final_scores=0.43282443,
final_local_scores=0.41078526, pem=0.0, similarity=16.348364]
10: 417429 Socialist_Party_of_Serbia
    [pem_pos=19, final_scores=0.37067074,
final_local_scores=0.40458208, pem=-4.5017786,
similarity=40.644623]
11: 29265 Serbia
    [pem_pos=0, final_scores=0.635001,
final_local_scores=0.6032531, pem=-0.34077588,
similarity=23.945183]
12: 417429 Socialist_Party_of_Serbia
    [pem_pos=19, final_scores=0.38302094,
final_local_scores=0.41807288, pem=-4.5017786,
similarity=41.047215]
```

```
13: 18998750 Reuters
    [pem_pos=0, final_scores=0.37685996,
final_local_scores=0.4564703, pem=0.0, similarity=17.711697]
14: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.38407224,
final_local_scores=0.32022297, pem=-0.7596219,
similarity=17.776747]
```

Sample: b'455_Croatian'

[1 1 Croatian 1] lending rate falls to 8.0 vs 9.1 pct . [2 2 ZAGREB 2] 1996-08-26 The [3 3 Croatian 3] lending rate fell to 8.0 percent on Monday from last Friday 's 9.1 percent after thin demand for kuna forced bank lenders to trim their rates further . Interbank call money was down to 8.0 from 10.0 percent . Insurers ' five-10 and 10-15 day loans were made at a steady 8.0 percent after last week 's drop from an earlier 15 percent . Total Zagreb Money Market settlements shrank to 17.4 million kuna , of which dealers put new borrowing at a meagre 5.5 million . Supply stood at a high 180 million kuna . Overnight trade at the weekend left a surplus of 1.2 billion kuna on the supply side after 40 million were settled . [4 4 Croatia 4] 's central bank again stayed out of the foreign exchange market . It calculated the kuna midrates for Tuesday 's trade stronger at 5.2420 against the dollar and slightly weaker at 3.5486 against the [5 5 German 5] mark . -- Kolumbina Bencevic , Zagreb Newsroom , 385-1-4557075

```
1: 5573 Croatia
2: 55906 Zagreb
3: 5573 Croatia
4: 5573 Croatia
5: 11867 Germany
1: 5573 Croatia
    [pem_pos=1, final_scores=0.51976293,
final_local_scores=0.44925928, pem=-1.1038951,
similarity=23.499659]
2: 55906 Zagreb
    [pem_pos=0, final_scores=0.47571123,
final_local_scores=0.43743777, pem=-0.06669851,
similarity=17.506445]
3: 5573 Croatia
    [pem_pos=1, final_scores=0.5337783,
final_local_scores=0.4645689, pem=-1.1038951,
similarity=23.95653]
4: 5573 Croatia
    [pem_pos=0, final_scores=0.63089436,
final_local_scores=0.57065344, pem=-0.35021794,
similarity=23.02369]
5: 11867 Germany
    [pem_pos=1, final_scores=0.82826906,
final_local_scores=0.8452148, pem=-1.0465823,
similarity=35.004086]
```

Sample: b'456_SVCD'

SVCD Bulgaria air traffic controllers to strike September 3 . [1 1 SOFIA 1] 1996-08-26 [2 2 Bulgarian 2] air traffic controllers will go on strike on September 3 demanding higher pay , the chief of the [3 3 Bulgarian 3] association of air traffic controllers (Bulatka) said on Monday . Stefan Raichev told a news conference the strike by more than half of the 1,380 traffic controllers and technicians would paralyse traffic which has increased by 10 percent since last year . More than 1,500 planes per day fly over [4 4 Bulgaria 4] , in a strategic location between [5 5 Europe 5] and the [6 6 Middle 6] and the Far East , Raichev said . The director general of the air traffic service Valentin Valkov said last Friday that a controllers ' strike would be illegal . Valkov said he could not curb the summer charter flights of national carrier [7 7 Balkan Airlines 7] , which carries thousands of foreign tourists to the [8 8 Bulgarian 8] [9 9 Black Sea 9] resorts . But Raichev said it would be " discrimination " against the other airlines if only Balkan planes were guided in . " Under the law before launching the strike we have to sign an agreement with our employer for a minimal transport servicing of emergency flights , " Raichev said . Raichev said a lock-out with military air controllers was impossible as they did not speak [10 10 English 10] . The controllers are demanding the monthly wage be increased to \$ 1,000 per month from the current \$ 230 , as well as the resignation of the air traffic service 's management . They also demand the financial separation of the 350 air controllers from the technical staff . -- Liliana Semerdjieva , Sofia Newsroom , 359-2-84561

1: 57644 Sofia

2: 3415 Bulgaria

3: 3415 Bulgaria

4: 3415 Bulgaria

5: 9239 Europe

6: 19323 Middle_East

7: 262416 Balkan_Bulgarian_Airlines

8: 3415 Bulgaria

9: 3386 Black_Sea

10: 8569916 English_language

1: 57644 Sofia

[pem_pos=0, final_scores=0.51247585, final_local_scores=0.49875236, pem=-0.10680699, similarity=19.55431]

2: 3415 Bulgaria

[pem_pos=1, final_scores=0.6352634, final_local_scores=0.5928822, pem=-1.1560022, similarity=28.06902]

3: 3415 Bulgaria

[pem_pos=1, final_scores=0.66765463, final_local_scores=0.6282647, pem=-1.1560022,

```
similarity=29.124905]
4: 3415 Bulgaria
    [pem_pos=0, final_scores=0.6854623,
final_local_scores=0.6477169, pem=-0.31727514,
similarity=25.144268]
5: 9239 Europe
    [pem_pos=0, final_scores=0.5897892,
final_local_scores=0.5786725, pem=-0.086338006,
similarity=21.827976]
6: 19323 Middle_East
    [pem_pos=2, final_scores=0.5027869,
final_local_scores=0.5138574, pem=-2.4908755,
similarity=32.970013]
7: 262416 Balkan_Bulgarian_Airlines
    [pem_pos=0, final_scores=0.44194144,
final_local_scores=0.5478782, pem=0.0, similarity=20.439491]
8: 3415 Bulgaria
    [pem_pos=1, final_scores=0.75299007,
final_local_scores=0.7214807, pem=-1.1560022,
similarity=31.906656]
9: 3386 Black_Sea
    [pem_pos=0, final_scores=0.40978307,
final_local_scores=0.4158063, pem=0.0, similarity=16.498201]
10: 8569916 English_language
    [pem_pos=0, final_scores=0.45386624,
final_local_scores=0.5081319, pem=-0.640266,
similarity=22.73525]
```

Sample: b'457_Yeltsin'

[1 1 Yeltsin 1] 's wife has kidney operation . [2 2 MOSCOW 2]
1996-08-26 Naina Yeltsin , wife of the [3 3 Russian 3]
president , has undergone " a planned operation " on her left
kidney and is in a satisfactory condition , [4 4 Itar-Tass 4]
news agency said on Monday . Tass quoted the [5 5 Kremlin 5]
press service as saying the operation took place on Saturday
in the [6 6 Central Clinical Hospital 6] which treats top
officials . Mrs Yeltsin would be released from hospital in a
few days . Doctor Sergei Mironov told Tass Naina was " in
permanent contact " with her husband and two daughters ,
Yelena and Tatyana . The state of health of [7 7 Boris Yeltsin
7] , who had two heart attacks last year , has been the
centre of media and market speculation after he won a second
term in office in the July 3 election run-off and all but
disappeared from the public eye . A presidential spokesman
said on Monday [8 8 Yeltsin 8] was in [9 9 Moscow 9] but
could give no details about his agenda or whether meetings
were planned . [10 10 Yeltsin 10] is expected to go on
holiday in the near future , but officials have not yet said
where he will go or when .

1: 23278520 Boris_Yeltsin
2: 19004 Moscow
3: 25391 Russia

```
4: 434852 Information_Telegraph_Agency_of_Russia
5: 381075 Moscow_Kremlin
6: 11101932 Central_Clinical_Hospital
7: 23278520 Boris_Yeltsin
8: 23278520 Boris_Yeltsin
9: 19004 Moscow
10: 23278520 Boris_Yeltsin
1: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.502525,
final_local_scores=0.45374382, pem=0.0, similarity=17.630333]
2: 19004 Moscow
    [pem_pos=0, final_scores=0.62653047,
final_local_scores=0.62236017, pem=-0.106582455,
similarity=23.241795]
3: 25391 Russia
    [pem_pos=1, final_scores=0.6807274,
final_local_scores=0.72236025, pem=-1.0515022,
similarity=31.364614]
4: 434852 Information_Telegraph_Agency_of_Russia
    [pem_pos=0, final_scores=0.7759636,
final_local_scores=0.85132813, pem=0.0, similarity=29.495043]
5: 381075 Moscow_Kremlin
    [pem_pos=1, final_scores=0.25829834,
final_local_scores=0.24848783, pem=-0.76193243,
similarity=15.648592]
6: 11101932 Central_Clinical_Hospital
    [pem_pos=0, final_scores=0.44676703,
final_local_scores=0.5290893, pem=0.0, similarity=19.878792]
7: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.62778527,
final_local_scores=0.59057164, pem=0.0, similarity=21.713549]
8: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.5220408,
final_local_scores=0.4750619, pem=0.0, similarity=18.266508]
9: 19004 Moscow
    [pem_pos=0, final_scores=0.661827,
final_local_scores=0.66091627, pem=-0.106582455,
similarity=24.392385]
10: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.6806944,
final_local_scores=0.6483668, pem=0.0, similarity=23.438272]
```

Sample: b'458_Mostostal'
Mostostal Z shareholders approve bonds . [1 1 WARSAW 1]
1996-08-26 Shareholders of the [2 2 Polish 2] construction
firm Mostostal Zabrze Holding SA approved a 25- million-
zloty , five-year convertible issue with a par value of 100
zlotys each , the company 's spokesman said on Monday . Piotr
Grabowski told [3 3 Reuters 3] that a [4 4 Mostostal 4]
extraordinary shareholders meeting on Saturday had decided
that the price of the 250,000 bonds would be set as a 10-
session average price of [5 5 Mostostal 5] shares plus a

premium . " The premium will be no lower than 15 percent and will be set by management before the issue ... which we expect at the beginning of next year , " Grabowski said . [6 6] Mostostal [7 7 Poland 7] , based in southern [7 7 Poland 7] , wants to offer the bonds to large investors in a public offering , paying an annual coupon of no more than 80 percent of the yield of the benchmark 52-week T-bill . The firm has signed an agreement with the Polish Development Bank (PBR) to manage the issue and plans them to be listed on the [8 8 Wars aw 8] bourse 's bond market . Grabowski also said shareholders approved the issue of 1.6 million shares for which the bonds could be exchanged , a ratio of 6.4 shares per bond . He also said shareholders approved the issue of 2.6 million new shares , two million of which are earmarked for large investors and 600,000 for retail domestic investors . The price will be determined through book-building . [9 9 Mostostal 9] plans to use the proceeds from the issues to add companies to its holding and modernise its plant . -- Warsaw Newsroom +48 22 653 9700

1: 32908 Warsaw
2: 22936 Poland
3: 18998750 Reuters
4: 2038354 Mostostal
5: 2038354 Mostostal
6: 2038354 Mostostal
7: 22936 Poland
8: 32908 Warsaw
9: 2038354 Mostostal
1: 32908 Warsaw
 [pem_pos=0, final_scores=0.47402948,
final_local_scores=0.5023279, pem=-0.14300999,
similarity=19.857891]
2: 22936 Poland
 [pem_pos=0, final_scores=0.555451,
final_local_scores=0.58213276, pem=-0.81518245,
similarity=25.894806]
3: 18998750 Reuters
 [pem_pos=0, final_scores=0.3359869,
final_local_scores=0.4539832, pem=0.0, similarity=17.637476]
4: 2038354 Mostostal
 [pem_pos=0, final_scores=0.46143085,
final_local_scores=0.4850335, pem=0.0, similarity=18.564081]
5: 2038354 Mostostal
 [pem_pos=0, final_scores=0.5445451,
final_local_scores=0.5758232, pem=0.0, similarity=21.273426]
6: 2038354 Mostostal
 [pem_pos=0, final_scores=0.44358033,
final_local_scores=0.4655345, pem=0.0, similarity=17.982191]
7: 22936 Poland
 [pem_pos=0, final_scores=0.46701688,
final_local_scores=0.485532, pem=-0.15333635,
similarity=19.412823]

```
8: 32908 Warsaw
    [pem_pos=0, final_scores=0.53779763,
final_local_scores=0.5719849, pem=-0.14300999,
similarity=21.936594]
9: 2038354 Mostostal
    [pem_pos=0, final_scores=0.46228403,
final_local_scores=0.4859655, pem=0.0, similarity=18.591892]
```

Sample: b'459_Unknown'

Unknown group kidnaps [1 1 Dutch 1] couple in [2 2 Costa Rica 2] . [3 3 SAN JOSE 3] , [4 4 Costa Rica 4] 1996-08-26

Kidnappers have seized a [5 5 Dutch 5] couple who manage a farm in northern [6 6 Costa Rica 6] and have demanded a \$ 1.5 million ransom , authorities said on Monday . Officials said Humberto Hueite Zyrecha and his wife Jetty Kors , both 50 , were kidnapped late on Saturday or early Sunday . Col . Misael Valerio , head of border police at the Security Ministry , told reporters surveillance of the northern border with [7 7 Nicaragua 7] has been stepped up to keep the kidnappers from fleeing [8 8 Costa Rica 8] . Preliminary reports said the couple were kidnapped on the " Altamira " farm , owned by [9 9 Dutchman 9] Richard Wisinga , at Aguas Zarcas de San Carlos near the Costa Rica-Nicaragua border . Unconfirmed news reports said a vehicle belonging to them was found abandoned at Santa Maria de Pocosol , about 12 miles (20 km) north of the site of the abduction . The reports said authorities found a statement , supposedly from the kidnappers , in the vehicle , addressed to Wisinga and demanding a \$ 1.5 million ransom . Wisinga was believed to be travelling in [10 10 Europe 10] , they said . A [11 11 Swiss 1 1] tourist guide and a [12 12 German 12] tourist were kidnapped in the same area on New Year 's Day by a group of former [13 13 Nicaraguan 13] guerrillas . Regula Susana Siegfried , 50 , and Nicola Fleuchaus , 25 , were released after 71 days after a \$ 200,000 ransom was paid . In a bizarre twist , Fleuchaus and the leader of the kidnappers , Julio Cesar Vega Rojas , developed a sentimental attachment during her captivity , according to photographs printed in [14 14 Costa Rican 14] newspapers and court documents including a love letter written by Vega .

```
1: 21148 Netherlands
2: 5551 Costa_Rica
3: 219665 San_José,_Costa_Rica
4: 5551 Costa_Rica
5: 21148 Netherlands
6: 5551 Costa_Rica
7: 21362 Nicaragua
8: 5551 Costa_Rica
9: 1656551 Dutch_people
10: 9239 Europe
11: 26748 Switzerland
12: 11867 Germany
```

```
13: 21362 Nicaragua
14: 5551 Costa_Rica
1: 21148 Netherlands
    [pem_pos=0, final_scores=0.5597045,
final_local_scores=0.5609117, pem=-0.7943452,
similarity=25.14821]
2: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.4514352,
final_local_scores=0.4109239, pem=-0.30652362,
similarity=18.019424]
3: 219665 San_José,_Costa_Rica
    [pem_pos=1, final_scores=0.58869,
final_local_scores=0.60720986, pem=-2.9430156,
similarity=38.21464]
4: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.58391625,
final_local_scores=0.5556393, pem=-0.30652362,
similarity=22.338022]
5: 21148 Netherlands
    [pem_pos=0, final_scores=0.61206377,
final_local_scores=0.6181063, pem=-0.7943452,
similarity=26.855011]
6: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.548583,
final_local_scores=0.5170431, pem=-0.30652362,
similarity=21.186234]
7: 21362 Nicaragua
    [pem_pos=0, final_scores=0.6314755,
final_local_scores=0.6166695, pem=-0.24729352,
similarity=23.837181]
8: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.54633325,
final_local_scores=0.51458555, pem=-0.30652362,
similarity=21.112896]
9: 1656551 Dutch_people
    [pem_pos=1, final_scores=0.21986295,
final_local_scores=0.24041599, pem=-1.4144723,
similarity=18.956326]
10: 9239 Europe
    [pem_pos=0, final_scores=0.6529379,
final_local_scores=0.6588162, pem=-0.086338006,
similarity=24.219624]
11: 26748 Switzerland
    [pem_pos=0, final_scores=0.7392839,
final_local_scores=0.7578353, pem=-0.1392413,
similarity=27.462246]
12: 11867 Germany
    [pem_pos=1, final_scores=0.74174494,
final_local_scores=0.7659647, pem=-1.0465823,
similarity=32.639107]
13: 21362 Nicaragua
    [pem_pos=0, final_scores=0.49203128,
```

```
final_local_scores=0.46434778, pem=0.0, similarity=17.946777]
14: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.5335391,
final_local_scores=0.50060993, pem=0.0, similarity=19.028912]
```

Sample: b'460_Interacciones'

Interacciones ups [1 1 Mexico 1] GDP forecast , lowers peso . [2 2 MEXICO CITY 2] 1996-08-26 Interacciones brokerage on Monday raised its forecasts for 1996 gross domestic product growth to 4.3 percent from 3.8 percent , but it kept its 1997 projection unchanged at 4.5 percent , a statement said . Economist Alonso Cervera said the revisions were chiefly fueled by stronger than expected growth of 7.2 percent in the second quarter and he forecast an annual GDP rise of 6.1 percent in the third quarter and 4.9 percent in Q4 . The firm also revised down its year-end peso forecast to 7.85-8.15 per dollar from 8.20-8.50 . It forecast an end 1997 peso in the range of 9.20-9.40 . Interacciones kept its 1996 and 1997 inflation forecast unchanged at 26 percent and 20 percent , expecting the government 's 1997 inflation target to be around 15 percent . It bumped up its average interest rate projection in 1997 to 25.8 percent from 23 percent . Fiscal policy was expected to be loosened a bit in the second half of this year , boosting growth without running into a deficit . " Next year , the fiscal policy will have less margin of freedom because the government will have to face commitments that it has taken on under different support schemes for the banks , debtors and firms , " the brokerage said . -- Henry Tricks , [3 3 Mexico City 3] newsroom +525 728-9560

```
1: 3966054 Mexico
2: 18987 Mexico_City
3: 18987 Mexico_City
1: 3966054 Mexico
    [pem_pos=0, final_scores=0.6708321,
final_local_scores=0.6802005, pem=-0.18258883,
similarity=25.381199]
2: 18987 Mexico_City
    [pem_pos=0, final_scores=0.48725468,
final_local_scores=0.42350465, pem=-0.028392298,
similarity=16.882338]
3: 18987 Mexico_City
    [pem_pos=0, final_scores=0.5378397,
final_local_scores=0.47876114, pem=-0.028392298,
similarity=18.531301]
```

Sample: b'461_Bancomext'

[1 1 Bancomext 1] official says [2 2 Mexico 2] peso level suitable . [3 3 MEXICO CITY 3] 1996-08-26 A top official of export development bank [4 4 Bancomext 4] said the peso 's appreciation against the dollar this year has not hurt demand for [5 5 Mexican 5] exports . " We have a suitable exchange

rate . We have n't felt a drop-off in demand We have no basis for saying that the exchange rate is affecting exports , " Rafael Moreno Turrent , deputy director general for foreign trade promotion at the bank , said at a news conference .

1: 21205898 Bancomext
2: 3966054 Mexico
3: 18987 Mexico_City
4: 21205898 Bancomext
5: 3966054 Mexico
1: 21205898 Bancomext
[pem_pos=0, final_scores=0.3476135,
final_local_scores=0.4192375, pem=0.0, similarity=16.600595]
2: 3966054 Mexico
[pem_pos=0, final_scores=0.5684106,
final_local_scores=0.58211964, pem=-0.18258883,
similarity=22.454271]
3: 18987 Mexico_City
[pem_pos=0, final_scores=0.37585264,
final_local_scores=0.41343218, pem=-0.028392298,
similarity=16.581757]
4: 21205898 Bancomext
[pem_pos=0, final_scores=0.4257949,
final_local_scores=0.5046388, pem=0.0, similarity=19.149141]
5: 3966054 Mexico
[pem_pos=0, final_scores=0.62148374,
final_local_scores=0.6400939, pem=-0.3968012,
similarity=25.349258]

Sample: b'462_Colombian'

[1 1 Colombian 1] peso closes lower , importers buy dtrs . [2
2 BOGOTA 2] 1996-08-26 [3 3 Colombia 3] 's peso closed lower at 1,044 after coming under pressure from dollar-buying by importers seeking to meet their commitments abroad , interbank dealers said . It was the second consecutive session that saw the peso , which ended at 1,041 Friday , close lower . " The (dollar 's) rise was due to demand from importers who have a lot of wire transfers accumulated , " one dealer said . Contributing to the peso 's weakness , another dealer said banks who might have sold greenbacks in Monday 's trading appeared to shy away from the market . This , according to the trader , was typical of month-end position-squaring -- and the fact that a stronger [4 4 U.S. 4] currency helps limit some foreign exchange losses . Dealers reported 355 trades for a total of \$ 142 million and said the peso hit an intra-day high of 1,037 before heading lower into the close . -- Juan Guillermo Londono , [5 5 Bogota 5] newsroom , 571 610 7944

1: 5222 Colombia
2: 211271 Bogotá
3: 5222 Colombia
4: 3434750 United_States

```
5: 211271 Bogotá
1: 5222 Colombia
    [pem_pos=0, final_scores=0.49688977,
final_local_scores=0.45334077, pem=-0.32036036,
similarity=19.360474]
2: 211271 Bogotá
    [pem_pos=0, final_scores=0.5960452,
final_local_scores=0.56742835, pem=0.0, similarity=21.022907]
3: 5222 Colombia
    [pem_pos=0, final_scores=0.5930777,
final_local_scores=0.55841154, pem=-0.22328302,
similarity=21.968075]
4: 3434750 United_States
    [pem_pos=0, final_scores=0.48451728,
final_local_scores=0.50086075, pem=-0.14075686,
similarity=19.801853]
5: 211271 Bogotá
    [pem_pos=0, final_scores=0.47070986,
final_local_scores=0.43051863, pem=0.0, similarity=16.937248]
```

Sample: b'463_Japan'

[1 1 Japan 1] 's [2 2 Hashimoto 2] leaves [3 3 Brazil 3] for
[4 4 Peru 4] . [5 5 BRASILIA 5] 1996-08-26 [6 6 Japanese 6]
Prime Minister [7 7 Ryutaro Hashimoto 7] left [8 8 Brasilia 8]
on Monday for [9 9 Lima 9] , the penultimate stop on a 10-day
[10 10 Latin American 10] tour , a [11 11 Brazilian 11]
Foreign Ministry spokeswoman said . [12 12 Hashimoto 12] ,
who has already visited [13 13 Mexico 13] and [14 14 Chile 14]
, spent three days in [15 15 Brazil 15] . After [16 16 Peru 16]
, he was due to go to [17 17 Costa Rica 17] .

```
1: 15573 Japan
2: 368549 Ryutaro_Hashimoto
3: 3383 Brazil
4: 170691 Peru
5: 4752 Brasília
6: 15573 Japan
7: 368549 Ryutaro_Hashimoto
8: 4752 Brasília
9: 85423 Lima
10: 18524 Latin_America
11: 3383 Brazil
12: 368549 Ryutaro_Hashimoto
13: 3966054 Mexico
14: 5489 Chile
15: 3383 Brazil
16: 170691 Peru
17: 5551 Costa_Rica
1: 15573 Japan
    [pem_pos=0, final_scores=0.7774914,
final_local_scores=0.79558355, pem=-0.13549899,
similarity=28.568377]
2: 368549 Ryutaro_Hashimoto
```

```
[pem_pos=0, final_scores=0.8307944,
final_local_scores=0.9493591, pem=0.0, similarity=32.420483]
3: 3383 Brazil
[pem_pos=0, final_scores=0.6027097,
final_local_scores=0.5885631, pem=-0.21313342,
similarity=22.812662]
4: 170691 Peru
[pem_pos=0, final_scores=0.63262385,
final_local_scores=0.62996596, pem=-0.23383401,
similarity=24.16078]
5: 4752 Brasília
[pem_pos=0, final_scores=0.43504733,
final_local_scores=0.433097, pem=0.0, similarity=17.01419]
6: 15573 Japan
[pem_pos=0, final_scores=0.974479,
final_local_scores=1.0107627, pem=-0.99076414,
similarity=39.64082]
7: 368549 Ryutaro_Hashimoto
[pem_pos=0, final_scores=0.8766349,
final_local_scores=0.98183626, pem=0.0, similarity=33.389664]
8: 4752 Brasília
[pem_pos=0, final_scores=0.34234887,
final_local_scores=0.331838, pem=0.0, similarity=13.992421]
9: 85423 Lima
[pem_pos=0, final_scores=0.5463231,
final_local_scores=0.5470033, pem=-0.32765645,
similarity=22.19523]
10: 18524 Latin_America
[pem_pos=0, final_scores=0.37446243,
final_local_scores=0.36218664, pem=-0.13736542,
similarity=15.645099]
11: 3383 Brazil
[pem_pos=0, final_scores=0.7081587,
final_local_scores=0.7037501, pem=-0.2733083,
similarity=26.57731]
12: 368549 Ryutaro_Hashimoto
[pem_pos=0, final_scores=0.90788263,
final_local_scores=1.0335662, pem=0.0, similarity=34.933388]
13: 3966054 Mexico
[pem_pos=0, final_scores=0.68438053,
final_local_scores=0.68326575, pem=-0.18258883,
similarity=25.472672]
14: 5489 Chile
[pem_pos=0, final_scores=0.6077896,
final_local_scores=0.59946024, pem=-0.27404457,
similarity=23.469099]
15: 3383 Brazil
[pem_pos=0, final_scores=0.6788508,
final_local_scores=0.6717357, pem=-0.21313342,
similarity=25.294699]
16: 170691 Peru
[pem_pos=0, final_scores=0.597956,
```

```
final_local_scores=0.59209657, pem=-0.23383401,
similarity=23.030682]
17: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.53046566,
final_local_scores=0.54247004, pem=-0.30652362,
similarity=21.945024]

Sample: b'464_Nicaraguan'
[1 1 Nicaraguan 1] president has operation in [2 2 U.S 2] . [
3 3 MANAGUA 3] , [4 4 Nicaragua 4] 1996-08-26 [5 5 Nicaraguan
5] President [6 6 Violeta Chamorro 6] underwent successful
surgery in the [7 7 United States 7] on Monday to correct a
compression in her lower spinal column , the government
said . Doctors at [8 8 Johns Hopkins Hospital 8] in [9 9 Balt
imore 9] found an inflammation in her spinal column that ,
once treated , will free her of chronic pain in her back and
one leg that has limited her movement , a government
statement said . " ( [10 10 Chamorro 10] ) is in good health
and , according to her doctors , will be able to return to [11
11 Nicaragua 11] next week , " it said . [12 12 Chamorro
12] , 66 , had complained of lower back pain since a trip to [
13 13 Taiwan 13] in May , when pain forced her to go to
Taipei University Hospital for an examination . She suffers
from osteoporosis , a disease that weakens the bones , and
has repeatedly flown to [14 14 Washington 14] for treatment .
She is scheduled to step down in January after a term of
nearly seven years .
1: 21362 Nicaragua
2: 3434750 United_States
3: 57042 Managua
4: 21362 Nicaragua
5: 21362 Nicaragua
6: 264238 Violeta_Chamorro
7: 3434750 United_States
8: 218873 Johns_Hopkins_Hospital
9: 26997138 Baltimore
10: 264238 Violeta_Chamorro
11: 21362 Nicaragua
12: 264238 Violeta_Chamorro
13: 25734 Taiwan
14: 108956 Washington,_D.C.
1: 21362 Nicaragua
    [pem_pos=0, final_scores=0.46313703,
final_local_scores=0.4304806, pem=0.0, similarity=16.936113]
2: 3434750 United_States
    [pem_pos=0, final_scores=0.412071,
final_local_scores=0.4012553, pem=0.0, similarity=16.063972]
3: 57042 Managua
    [pem_pos=0, final_scores=0.4371357,
final_local_scores=0.42301953, pem=0.0, similarity=16.713459]
4: 21362 Nicaragua
    [pem_pos=0, final_scores=0.5407235,
```

```
final_local_scores=0.515232, pem=-0.24729352,
similarity=20.810085]
5: 21362 Nicaragua
    [pem_pos=0, final_scores=0.46360493,
final_local_scores=0.4309917, pem=0.0, similarity=16.951365]
6: 264238 Violeta_Chamorro
    [pem_pos=0, final_scores=0.8021653,
final_local_scores=0.83780146, pem=0.0, similarity=29.09138]
7: 3434750 United_States
    [pem_pos=0, final_scores=0.5183768,
final_local_scores=0.5173783, pem=-0.073457584,
similarity=19.928787]
8: 218873 Johns_Hopkins_Hospital
    [pem_pos=0, final_scores=0.32064074,
final_local_scores=0.42397338, pem=0.0, similarity=16.741924]
9: 26997138 Baltimore
    [pem_pos=0, final_scores=0.44425625,
final_local_scores=0.47831285, pem=-0.19949137,
similarity=19.448387]
10: 264238 Violeta_Chamorro
    [pem_pos=0, final_scores=0.86327773,
final_local_scores=0.91816765, pem=0.0, similarity=31.489666]
11: 21362 Nicaragua
    [pem_pos=0, final_scores=0.5167063,
final_local_scores=0.48899686, pem=-0.24729352,
similarity=20.027174]
12: 264238 Violeta_Chamorro
    [pem_pos=0, final_scores=0.8808642,
final_local_scores=0.9373782, pem=0.0, similarity=32.06295]
13: 25734 Taiwan
    [pem_pos=0, final_scores=0.34681815,
final_local_scores=0.36942518, pem=-0.077886656,
similarity=15.537657]
14: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.49603873,
final_local_scores=0.486678, pem=-1.7496481,
similarity=28.128014]
```

Sample: b'465_Eight'

Eight drown in [1 1 Venezuelan 1] boating accident . [2 2 MA RACAIBO 2] , [3 3 Venezuela 3] 1996-08-26 Eight members of a family , five of them children aged between two and seven , drowned when their small boat sank on [4 4 Lake Maracaibo 4] in western [5 5 Venezuela 5] early on Monday , authorities said . The accident happened when the Sanchez Zarraga family took their boat out for a nighttime spin , Civil Defence and Coast Guard officials said . The cause of the sinking was not known but officials said the boat had a hole in the stern and no lifejackets . Three members of the party were rescued unhurt .

```
1: 32374 Venezuela
2: 249847 Maracaibo
```

```
3: 32374 Venezuela
4: 265839 Lake_Maracaibo
5: 32374 Venezuela
1: 32374 Venezuela
    [pem_pos=0, final_scores=0.63921523,
final_local_scores=0.59137744, pem=-0.11529429,
similarity=22.364584]
2: 249847 Maracaibo
    [pem_pos=0, final_scores=0.47443444,
final_local_scores=0.49105054, pem=0.0, similarity=18.74364]
3: 32374 Venezuela
    [pem_pos=0, final_scores=0.62231565,
final_local_scores=0.57291716, pem=-0.21226841,
similarity=22.341053]
4: 265839 Lake_Maracaibo
    [pem_pos=0, final_scores=0.38716722,
final_local_scores=0.4404118, pem=0.0, similarity=17.23248]
5: 32374 Venezuela
    [pem_pos=0, final_scores=0.5853974,
final_local_scores=0.5325897, pem=-0.21226841,
similarity=21.1376]
```

Sample: b'466_Banco'

```
[1 1 Banco de Mexico 1] to inject 3.412 bln pesos . [2 2 MEX
ICO CITY 2] 1996-08-26 [3 3 Banco de Mexico 3] sought to
inject 3.412 billion pesos in the secondary market on Monday
through three credit auctions , dealers said . The auctions
were offered as follows : AMOUNT TERM 1.206 bln 9 days 1.206
bln 3 days 1.000 bln 1 day - [4 4 Mexico City 4] newsroom 525
728-9559
1: 180390 Bank_of_Mexico
2: 18987 Mexico_City
3: 180390 Bank_of_Mexico
4: 18987 Mexico_City
1: 180390 Bank_of_Mexico
    [pem_pos=0, final_scores=0.5073739,
final_local_scores=0.5114201, pem=0.0, similarity=19.351507]
2: 18987 Mexico_City
    [pem_pos=0, final_scores=0.44050556,
final_local_scores=0.4383765, pem=-0.028392298,
similarity=17.326143]
3: 180390 Bank_of_Mexico
    [pem_pos=0, final_scores=0.5175047,
final_local_scores=0.5224864, pem=0.0, similarity=19.68175]
4: 18987 Mexico_City
    [pem_pos=0, final_scores=0.4589252,
final_local_scores=0.45849717, pem=-0.028392298,
similarity=17.926584]
```

Sample: b'467_Dutch'

```
[1 1 Dutch 1] couple kidnapped in [2 2 Costa Rica 2] . [3 3 S
AN JOSE 3] , [4 4 Costa Rica 4] 1996-08-26 A [5 5 Dutch 5]
```

couple who manage a farm in northern [6 6 Costa Rica 6] have been kidnapped by an unidentified group demanding \$ 1.5 million in ransom , authorities said on Monday . Officials said Humberto Hueite Zyrecha and his wife Jetty Kors , both 50 , were kidnapped late Saturday or early Sunday . Col . Misael Valerio , head of border police at the Security Ministry , told reporters that surveillance of the country 's northern border with [7 7 Nicaragua 7] has been stepped up to stop the kidnappers from fleeing [8 8 Costa Rica 8] . No other details were immediately available .

```
1: 21148 Netherlands
2: 5551 Costa_Rica
3: 219665 San_José,_Costa_Rica
4: 5551 Costa_Rica
5: 21148 Netherlands
6: 5551 Costa_Rica
7: 21362 Nicaragua
8: 5551 Costa_Rica
1: 21148 Netherlands
    [pem_pos=0, final_scores=0.54921025,
final_local_scores=0.5839169, pem=-0.7943452,
similarity=25.83473]
2: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.53370416,
final_local_scores=0.4858328, pem=-0.30652362,
similarity=20.254856]
3: 219665 San_José,_Costa_Rica
    [pem_pos=1, final_scores=0.58764136,
final_local_scores=0.6073707, pem=-2.9430156,
similarity=38.21944]
4: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.60777074,
final_local_scores=0.56673926, pem=-0.30652362,
similarity=22.669266]
5: 21148 Netherlands
    [pem_pos=0, final_scores=0.56051594,
final_local_scores=0.5962666, pem=-0.7943452,
similarity=26.203272]
6: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.5774944,
final_local_scores=0.53366697, pem=-0.30652362,
similarity=21.682323]
7: 21362 Nicaragua
    [pem_pos=0, final_scores=0.62302,
final_local_scores=0.61183095, pem=-0.24729352,
similarity=23.69279]
8: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.60281986,
final_local_scores=0.5613312, pem=-0.30652362,
similarity=22.50788]
```

Sample: b'468_RTRS'

RTRS - [1 1 Melbourne 1] train collides with truck , 15 injured . [2 2 MELBOURNE 2] 1996-08-26 Fifteen people were injured when a suburban passenger train and a truck collided at a street-level rail crossing in the [3 3 Australian 3] city of [4 4 Melbourne 4] on Monday , said rail and ambulance officials . The injured , which included a pregnant woman , were taken to hospital , but were not in a serious condition , an ambulance spokesman told [5 5 Reuters 5] . A [6 6 Public Transport Corporation 6] spokesman said the train , collided with the truck loaded with concrete pylons in the north [7 7 Melbourne 7] suburb of [8 8 Preston 8] just before 8.30 a.m .. The train was derailed when a corner of the driver 's compartment caught the rear of the truck , the spokesman said . The truck was overturned , spilling its load onto the crossing , and careered into the nearby Bell St Station . " Remarkably both the train driver and the truck driver were not injured , " the rail spokesman said . " It had the potential to be quite a nasty accident . " -- Sydney Newsroom 61-2 373-1800

1: 17306237 Melbourne
2: 17306237 Melbourne
3: 4689264 Australia
4: 17306237 Melbourne
5: 18998750 Reuters
6: 12927156 Public_Transport_Corporation
7: 17306237 Melbourne
8: 88918 Preston,_Lancashire
1: 17306237 Melbourne
[pem_pos=0, final_scores=0.47100002,
final_local_scores=0.40312606, pem=-0.16841523,
similarity=17.035667]
2: 17306237 Melbourne
[pem_pos=0, final_scores=0.5766521,
final_local_scores=0.51853484, pem=-0.16841523,
similarity=20.479696]
3: 4689264 Australia
[pem_pos=0, final_scores=0.6886638,
final_local_scores=0.7011671, pem=-0.491009,
similarity=27.68412]
4: 17306237 Melbourne
[pem_pos=0, final_scores=0.521823,
final_local_scores=0.45864242, pem=-0.16841523,
similarity=18.692387]
5: 18998750 Reuters
[pem_pos=0, final_scores=0.33483934,
final_local_scores=0.42416513, pem=0.0, similarity=16.747646]
6: 12927156 Public_Transport_Corporation
[pem_pos=0, final_scores=0.41160643,
final_local_scores=0.51569843, pem=0.0, similarity=19.479183]
7: 17306237 Melbourne
[pem_pos=0, final_scores=0.4977246,
final_local_scores=0.4323185, pem=-0.16841523,

```
similarity=17.906828]
8: 88918 Preston,_Lancashire
    [pem_pos=0, final_scores=0.39896774,
final_local_scores=0.41030675, pem=-1.5624295,
similarity=24.83082]

Sample: b'469_Police'
Police capture [1 1 Auckland 1] gunman . [2 2 AUCKLAND 2]
1996-08-27 A gunman being hunted on [3 3 Auckland 3] 's [4 4 N
orth Shore 4] was captured by police just after 9 a.m. on
Tuesday , [5 5 New Zealand Press Association 5] reported .
Senior Sergeant Dave Pearson of [6 6 Auckland 6] police said
John Grant Fagan was in police custody , but further details
of his capture would not be released until a press conference
later in the day . Fagan had earlier telephoned an [7 7 Auckl
and 7] radio station in a distraught state , saying he feared
for his life . Police said he did not have a weapon when
taken into custody and was now cooperating with them . [8 8 Ra
dio New Zealand 8] reported earlier that an armed man on
Monday entered the Northcote College swimming pool changing
sheds and confronted a 16-year-old schoolgirl . A shot was
fired , but onlookers managed to disarm the man . -- [9 9 Wel
lington 9] newsroom 64 4 473-4746
1: 18660332 Auckland
2: 18660332 Auckland
3: 18660332 Auckland
4: 350800 North_Shore_City
5: 3185516 New_Zealand_Press_Association
6: 18660332 Auckland
7: 18660332 Auckland
8: 384076 Radio_New_Zealand
9: 33804 Wellington
1: 18660332 Auckland
    [pem_pos=0, final_scores=0.57705224,
final_local_scores=0.51543045, pem=-0.297032,
similarity=21.086493]
2: 18660332 Auckland
    [pem_pos=0, final_scores=0.55862564,
final_local_scores=0.49530214, pem=-0.297032,
similarity=20.485823]
3: 18660332 Auckland
    [pem_pos=0, final_scores=0.54308325,
final_local_scores=0.47832447, pem=-0.297032,
similarity=19.979176]
4: 350800 North_Shore_City
    [pem_pos=9, final_scores=0.31777948,
final_local_scores=0.31658092, pem=-2.6431773,
similarity=27.911129]
5: 3185516 New_Zealand_Press_Association
    [pem_pos=0, final_scores=0.55362827,
final_local_scores=0.6644072, pem=0.0, similarity=23.916948]
6: 18660332 Auckland
```

```
[pem_pos=0, final_scores=0.48787183,
final_local_scores=0.4180144, pem=-0.297032,
similarity=18.179403]
7: 18660332 Auckland
    [pem_pos=0, final_scores=0.500687,
final_local_scores=0.43201303, pem=-0.297032,
similarity=18.59715]
8: 384076 Radio_New_Zealand
    [pem_pos=0, final_scores=0.4596871,
final_local_scores=0.49294674, pem=0.0, similarity=18.800226]
9: 33804 Wellington
    [pem_pos=0, final_scores=0.5317691,
final_local_scores=0.4787386, pem=-0.9265605,
similarity=23.415009]
```

Sample: b'470_Police'

Police seek fugitive after [1 1 Auckland 1] gun incident . [2 2 WELLINGTON 2] 1996-08-27 [3 3 Auckland 3] police were seeking an escaped gunman on Tuesday after an incident in which he fired at a 16-year-old schoolgirl , [4 4 Radio New Zealand 4] said . The man entered the Northcote College swimming pool changing sheds on Monday and told the girl and a friend : " You 're for it now . " A shot was fired , but onlookers managed to seize the gun , the radio said . Police , who were unsuccessful in finding the man overnight , described him as disturbed and dangerous . -- [5 5 Wellington 5] newsroom 64 4 473-4746

```
1: 18660332 Auckland
2: 33804 Wellington
3: 18660332 Auckland
4: 384076 Radio_New_Zealand
5: 33804 Wellington
1: 18660332 Auckland
    [pem_pos=0, final_scores=0.608354,
final_local_scores=0.5529074, pem=-0.297032,
similarity=22.20488]
2: 33804 Wellington
    [pem_pos=0, final_scores=0.50090206,
final_local_scores=0.4370836, pem=-0.9265605,
similarity=22.17194]
3: 18660332 Auckland
    [pem_pos=0, final_scores=0.5002009,
final_local_scores=0.43476665, pem=-0.297032,
similarity=18.679323]
4: 384076 Radio_New_Zealand
    [pem_pos=0, final_scores=0.5580618,
final_local_scores=0.61198443, pem=0.0, similarity=22.352549]
5: 33804 Wellington
    [pem_pos=0, final_scores=0.54620343,
final_local_scores=0.48656845, pem=-0.9265605,
similarity=23.648666]
```

Sample: b'471_SIMEX'

SIMEX Brent closed on Monday owing to 1 IPE holiday . [1 1 SI
NGAPORE 1] 1996-08-26 The Brent crude futures market on the [2
2 Singapore International Monetary Exchange 2] ([3 3 SIMEX 3]
) was closed on Monday in respect for a [4 4 U.K. 4] national
holiday . THE SIMEX Brent market keeps to the trading
schedule of the [5 5 International Petroleum Exchange 5] (2 I
PE) in [6 6 London 6] , which is closed for a [7 7 British 7]
bank holiday . Contracts traded in [8 8 Singapore 8] are
mutually offset against contracts traded in [9 9 London 9] .
-- Singapore Newsroom (+65 870 3081)

1: 27318 Singapore
2: 840431 Singapore_International_Monetary_Exchange
3: 840431 Singapore_International_Monetary_Exchange
4: 31717 United_Kingdom
5: 1049625 International_Petroleum_Exchange
6: 17867 London
7: 31717 United_Kingdom
8: 27318 Singapore
9: 17867 London
1: 1049625 International_Petroleum_Exchange
[winner: 51578 Tabebuia, pem_pos=1,
final_scores=-0.31470895, final_local_scores=-0.19479674,
pem=-0.3115333, similarity=-0.029243708]
RECALL MISS: 1049625 International_Petroleum_Exchange]
2: 1049625 International_Petroleum_Exchange
[winner: 11446063 Ipe_(program), pem_pos=0,
final_scores=-0.3515126, final_local_scores=-0.20195824,
pem=0.0, similarity=-1.9371237]
RECALL MISS: 1049625 International_Petroleum_Exchange]
1: 27318 Singapore
[pem_pos=0, final_scores=0.53555846,
final_local_scores=0.51145697, pem=-0.1922004,
similarity=20.397825]
2: 840431 Singapore_International_Monetary_Exchange
[pem_pos=0, final_scores=0.40249783,
final_local_scores=0.50904894, pem=0.0, similarity=19.280748]
3: 840431 Singapore_International_Monetary_Exchange
[pem_pos=0, final_scores=0.37770468,
final_local_scores=0.48196626, pem=0.0, similarity=18.472548]
4: 31717 United_Kingdom
[pem_pos=0, final_scores=0.31652868,
final_local_scores=0.28113514, pem=-0.13569236,
similarity=13.217262]
5: 1049625 International_Petroleum_Exchange
[pem_pos=0, final_scores=0.40671402,
final_local_scores=0.5391906, pem=0.0, similarity=20.180235]
6: 17867 London
[pem_pos=0, final_scores=0.547901,
final_local_scores=0.5350812, pem=-0.09086737,
similarity=20.551754]
7: 31717 United_Kingdom

```
[pem_pos=0, final_scores=0.4242776,
final_local_scores=0.39883447, pem=-0.48936236,
similarity=18.652958]
8: 27318 Singapore
[pem_pos=0, final_scores=0.4280817,
final_local_scores=0.3940549, pem=-0.1922004,
similarity=16.894312]
9: 17867 London
[pem_pos=0, final_scores=0.5192072,
final_local_scores=0.5037376, pem=-0.09086737,
similarity=19.616398]
```

Sample: b'472_Metro'

Metro slides 3.3 pct after market opens . [1 1 SINGAPORE 1]
1996-08-26 Shares in retailer Metro Holdings dropped 3.31
percent , or [2 2 Singapore 2] \$ 0.20 , to S\$ 5.85 minutes
after the market opened on Monday . By 0120 [3 3 GMT 3] ,
357,000 Metro shares had been traded . On Friday , Metro
Holdings topped gainers , soaring by S\$ 1.55 to close at S\$
6.05 on market rumours of a takeover bid by First Capital
Corp . The company said it was not aware of any reason for
the surge . -- [4 4 Singapore 4] newsroom (65 8703080)

```
1: 27318 Singapore
2: 27318 Singapore
3: 12701 Greenwich_Mean_Time
4: 27318 Singapore
1: 27318 Singapore
[pem_pos=0, final_scores=0.6061491,
final_local_scores=0.55417156, pem=-0.1922004,
similarity=21.672514]
2: 27318 Singapore
[pem_pos=0, final_scores=0.44183868,
final_local_scores=0.37468743, pem=-0.1922004,
similarity=16.316349]
3: 12701 Greenwich_Mean_Time
[pem_pos=0, final_scores=0.68654996,
final_local_scores=0.8043294, pem=0.0, similarity=28.092506]
4: 27318 Singapore
[pem_pos=0, final_scores=0.54188997,
final_local_scores=0.48397827, pem=-0.1922004,
similarity=19.577806]
```

Sample: b'473_Indonesian'

[1 1 Indonesian 1] rupiah stable in quiet late trading . [2 2
JAKARTA 2] 1996-08-26 The [3 3 Indonesian 3] rupiah was
stable against the dollar in quiet trading on Monday ,
dealers said . They said volume was thin following a public
holiday in [4 4 Hong Kong 4] and the [5 5 United Kingdom 5] .
" We did n't see anything from [6 6 Singapore 6] operators
either . It 's a pretty quiet day , " one foreign bank dealer
said . Spot rupiah was quoted at 2,342.0 / 42.5 at 0915 [7 7 G
MT 7] , unchanged from the opening level . It was softer in

the morning due to relatively ample rupiah liquidity but recovered later . Tomorrow and today rupiah closed at 2,342.00 / 42.45 and 2,341.5 / 42.0 , respectively . Another dealer said operators were reluctant to unload rupiah despite ample conditions due to the month-end factor . " There are two factors which determine the market at present . The liquidity at month-end and the next court hearing scheduled for Thursday , " the dealer said . [8 8 Megawati Sukarnoputri 8] , deposed leader of the [9 9 Indonesian Democratic Party 9] ([10 10 PDI 10]) has sued the government for ousting her as [11 11 PDI 11] leader . The central [12 12 Jakarta 12] court adjourned the case last Thursday for possible out-of-court settlement . The hearing is due to resume on Thursday . [13 13 Megawati 13] 's lawyers said they were still discussing a possible out of court settlement but they was not optimistic an agreement could be reached . The dealer said this issue remained a factor but its importance to the market seemed to diminish . He said liquidity would be the main concern for the next few days . Overnight swap was at 0.45 / 0.48 and tom / next at 0.50 / 0.55 . One-month swap was at 18.0 / 18.5 , two at 34.0 / 35.0 , three at 52.75 / 53.50 and six at 106.5 / 107.0 points . The central bank kept its intervention rate at 2,337 / 2,455 and the conversion rate at 2,337 / 2,383 on Monday .

1: 14579 Indonesia
2: 16275 Jakarta
3: 14579 Indonesia
4: 13404 Hong_Kong
5: 31717 United_Kingdom
6: 27318 Singapore
7: 12701 Greenwich_Mean_Time
8: 237564 Megawati_Sukarnoputri
9: 1010736 Indonesian_Democratic_Party
10: 1010736 Indonesian_Democratic_Party
11: 1010736 Indonesian_Democratic_Party
12: 16275 Jakarta
13: 237564 Megawati_Sukarnoputri
1: 14579 Indonesia
[pem_pos=1, final_scores=0.49253917,
final_local_scores=0.47948396, pem=-0.74208045,
similarity=22.43402]
2: 16275 Jakarta
[pem_pos=0, final_scores=0.4457491,
final_local_scores=0.41275418, pem=0.0, similarity=16.407122]
3: 14579 Indonesia
[pem_pos=1, final_scores=0.46091014,
final_local_scores=0.44493407, pem=-0.74208045,
similarity=21.402983]
4: 13404 Hong_Kong
[pem_pos=0, final_scores=0.4373541,
final_local_scores=0.4324535, pem=-0.22402541,
similarity=18.213274]

```
5: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.45911288,
final_local_scores=0.49213898, pem=-0.09933331,
similarity=19.316313]
6: 27318 Singapore
    [pem_pos=0, final_scores=0.5022375,
final_local_scores=0.48621404, pem=-0.1922004,
similarity=19.644526]
7: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.62137324,
final_local_scores=0.7421765, pem=0.0, similarity=26.23774]
8: 237564 Megawati_Sukarnoputri
    [pem_pos=0, final_scores=0.5258977,
final_local_scores=0.5415418, pem=0.0, similarity=20.2504]
9: 1010736 Indonesian_Democratic_Party
    [pem_pos=0, final_scores=0.4574595,
final_local_scores=0.51064754, pem=0.0, similarity=19.328453]
10: 1010736 Indonesian_Democratic_Party
    [pem_pos=2, final_scores=0.5781209,
final_local_scores=0.6424518, pem=-1.1770838,
similarity=29.662922]
11: 1010736 Indonesian_Democratic_Party
    [pem_pos=2, final_scores=0.6442527,
final_local_scores=0.7146908, pem=-1.1770838,
similarity=31.818678]
12: 16275 Jakarta
    [pem_pos=0, final_scores=0.4644314,
final_local_scores=0.4331618, pem=0.0, similarity=17.016125]
13: 237564 Megawati_Sukarnoputri
    [pem_pos=0, final_scores=0.3914656,
final_local_scores=0.39469522, pem=0.0, similarity=15.868206]
```

Sample: b'474_Senate'

1 Senate intelligence chairman in [1 1 Saudi 1] bomb probe .
[2 2 DUBAI 2] 1996-08-26 [3 3 U.S. Senate Intelligence Commit
tee 3] chairman [4 4 Arlen Specter 4] , who has questioned
whether Defence Secretary [5 5 William Perry 5] should resign
over the latest bombing in [6 6 Saudi Arabia 6] , met [7 7 Sa
udi 7] officials on Monday during a brief visit to the
kingdom . A [8 8 U.S. 8] embassy spokesman in [9 9 Riyadh 9]
said [10 10 Specter 10] , who arrived from neighbouring [11 11
Oman 11] on Sunday and left on Monday , had talks with [12 12
Saudi 12] and [13 13 American 13] officials in [14 14 Dhahran
14] , where 19 [15 15 U.S. 15] airmen were killed by a fuel
truck bomb on June 25 , and [16 16 Riyadh 16] . [17 17 Specter
17] met Crown [18 18 Prince Abdullah 18] and Minister of
Defence and Aviation Prince [19 19 Sultan 19] in [20 20 Jeddah
20] , [21 21 Saudi 21] state television and the official [22 2
2 Saudi Press Agency 22] reported . He had earlier visited [23
23 Japan 23] , [24 24 South Korea 24] and [25 25 China 25] . [
26 26 Specter 26] said after the bombing there should be a
shake-up at the [27 27 Pentagon 27] and questioned whether [28

[28 Perry 28] should resign . He said he was not satisfied with some of [29 29 Perry 29] 's answers to the committee 's questions in closed testimony last month . The question of whether [30 30 Perry 30] should resign remained open , the [31 31 Pennsylvania 31] Republican said . [32 32 FBI 32] Director [33 33 Louis Freeh 33] , who has twice visited [34 34 Saudi Arabia 34] to seek improved cooperation with [35 35 Saudi 35] investigators , told the committee he was not entirely satisfied with [36 36 Saudi 36] cooperation on the [37 37 Dha hran 37] bomb and a previous bomb attack in [38 38 Riyadh 38] . " If we 're to stay in [39 39 Saudi Arabia 39] , we need to have total cooperation , " [40 40 Specter 40] said . The [41 41 United States 41] has 5,000 [42 42 U.S. 42] air force and other military personnel in [43 43 Saudi Arabia 43] .

1: 349303 Saudi_Arabia

2: 211583 Dubai

3: 801696

United_States_Senate_Select_Committee_on_Intelligence

4: 367801 Arlen_Specter

5: 152783 William_Perry

6: 349303 Saudi_Arabia

7: 349303 Saudi_Arabia

8: 3434750 United_States

9: 51554 Riyadh

10: 367801 Arlen_Specter

11: 22316 Oman

12: 349303 Saudi_Arabia

13: 3434750 United_States

14: 6838977 Dhahran

15: 3434750 United_States

16: 51554 Riyadh

17: 367801 Arlen_Specter

18: 19186951 Abdullah_of_Saudi_Arabia

19: 403248 Sultan_bin_Abdulaziz

20: 277718 Jeddah

21: 349303 Saudi_Arabia

22: 2902711 Saudi_Press_Agency

23: 15573 Japan

24: 27019 South_Korea

25: 5405 China

26: 367801 Arlen_Specter

27: 20740978 The_Pentagon

28: 152783 William_Perry

29: 152783 William_Perry

30: 152783 William_Perry

31: 23332 Pennsylvania

32: 11127 Federal_Bureau_of_Investigation

33: 21810686 Louis_Freeh

34: 349303 Saudi_Arabia

35: 349303 Saudi_Arabia

36: 349303 Saudi_Arabia

37: 6838977 Dhahran

38: 51554 Riyadh
39: 349303 Saudi_Arabia
40: 367801 Arlen_Specter
41: 3434750 United_States
42: 3434750 United_States
43: 349303 Saudi_Arabia

1: 515101 United_States_Senate_Committee_on_Armed_Services
[winner: 24909346 United_States_Senate, pem_pos=0,
final_scores=-0.41898954, final_local_scores=-0.47826183,
pem=-1.2752931, similarity=-3.2473314]
RECALL MISS: 515101

United_States_Senate_Committee_on_Armed_Services]

1: 349303 Saudi_Arabia
[pem_pos=0, final_scores=0.511808,
final_local_scores=0.47650218, pem=0.0, similarity=18.309488]

2: 211583 Dubai
[pem_pos=0, final_scores=0.52618486,
final_local_scores=0.540799, pem=-0.08179128,
similarity=20.673029]

3: 801696

United_States_Senate_Select_Committee_on_Intelligence
[pem_pos=0, final_scores=0.26532412,
final_local_scores=0.29884386, pem=0.0, similarity=13.00781]

4: 367801 Arlen_Specter
[pem_pos=0, final_scores=0.6698115,
final_local_scores=0.72191995, pem=0.0, similarity=25.633244]

5: 152783 William_Perry
[pem_pos=0, final_scores=0.5069944,
final_local_scores=0.5546234, pem=-0.24904563,
similarity=21.99513]

6: 349303 Saudi_Arabia
[pem_pos=0, final_scores=0.4705693,
final_local_scores=0.4314552, pem=-0.25454172,
similarity=18.349436]

7: 349303 Saudi_Arabia
[pem_pos=0, final_scores=0.61056894,
final_local_scores=0.5843836, pem=0.0, similarity=21.528885]

8: 3434750 United_States
[pem_pos=0, final_scores=0.5082045,
final_local_scores=0.49934036, pem=-0.14075686,
similarity=19.756481]

9: 51554 Riyadh
[pem_pos=0, final_scores=0.34605068,
final_local_scores=0.3185722, pem=0.0, similarity=13.596542]

10: 367801 Arlen_Specter
[pem_pos=0, final_scores=0.40813953,
final_local_scores=0.44189727, pem=0.0, similarity=17.27681]

11: 22316 Oman
[pem_pos=0, final_scores=0.43372244,
final_local_scores=0.43215865, pem=-0.38991782,
similarity=19.106623]

12: 349303 Saudi_Arabia

[pem_pos=0, final_scores=0.6448987,
final_local_scores=0.6218837, pem=0.0, similarity=22.647963]
13: 3434750 United_States
[pem_pos=0, final_scores=0.5600148,
final_local_scores=0.55593514, pem=-0.12994486,
similarity=21.386587]
14: 6838977 Dhahran
[pem_pos=0, final_scores=0.37486166,
final_local_scores=0.3730923, pem=0.0, similarity=15.223532]
15: 3434750 United_States
[pem_pos=0, final_scores=0.5070683,
final_local_scores=0.49809915, pem=-0.14075686,
similarity=19.719444]
16: 51554 Riyadh
[pem_pos=0, final_scores=0.36045128,
final_local_scores=0.3343027, pem=0.0, similarity=14.065972]
17: 367801 Arlen_Specter
[pem_pos=0, final_scores=0.43085968,
final_local_scores=0.46671557, pem=0.0, similarity=18.017437]
18: 19186951 Abdullah_of_Saudi_Arabia
[pem_pos=0, final_scores=0.3158806,
final_local_scores=0.3147537, pem=-0.18784133,
similarity=14.504102]
19: 403248 Sultan_bin_Abdulaziz
[pem_pos=4, final_scores=0.49408203,
final_local_scores=0.54278344, pem=-3.752026,
similarity=40.691555]
20: 277718 Jeddah
[pem_pos=0, final_scores=0.33817595,
final_local_scores=0.31987026, pem=0.0, similarity=13.63528]
21: 349303 Saudi_Arabia
[pem_pos=0, final_scores=0.5992219,
final_local_scores=0.57198864, pem=0.0, similarity=21.158995]
22: 2902711 Saudi_Press_Agency
[pem_pos=0, final_scores=0.49110776,
final_local_scores=0.61676186, pem=0.0, similarity=22.495117]
23: 15573 Japan
[pem_pos=0, final_scores=0.5491091,
final_local_scores=0.6045363, pem=-0.13549899,
similarity=22.867146]
24: 27019 South_Korea
[pem_pos=0, final_scores=0.37947685,
final_local_scores=0.40868753, pem=-0.22449642,
similarity=17.50661]
25: 5405 China
[pem_pos=0, final_scores=0.46923727,
final_local_scores=0.49535793, pem=-0.15669827,
similarity=19.72433]
26: 367801 Arlen_Specter
[pem_pos=0, final_scores=0.4256956,
final_local_scores=0.4610746, pem=0.0, similarity=17.849098]
27: 20740978 The_Pentagon

```
[pem_pos=0, final_scores=0.30194265,
final_local_scores=0.28644854, pem=-0.14635207,
similarity=13.433794]
28: 152783 William_Perry
    [pem_pos=0, final_scores=0.5431624,
final_local_scores=0.5941314, pem=-0.24904563,
similarity=23.174128]
29: 152783 William_Perry
    [pem_pos=0, final_scores=0.49185145,
final_local_scores=0.538082, pem=-0.24904563,
similarity=21.501501]
30: 152783 William_Perry
    [pem_pos=0, final_scores=0.56280667,
final_local_scores=0.61558986, pem=-0.24904563,
similarity=23.814491]
31: 23332 Pennsylvania
    [pem_pos=0, final_scores=0.63902885,
final_local_scores=0.6830389, pem=-0.11810702,
similarity=25.11524]
32: 11127 Federal_Bureau_of_Investigation
    [pem_pos=0, final_scores=0.45106703,
final_local_scores=0.47622812, pem=0.0, similarity=18.30131]
33: 21810686 Louis_Freeh
    [pem_pos=0, final_scores=0.6544124,
final_local_scores=0.72469276, pem=0.0, similarity=25.71599]
34: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.45705217,
final_local_scores=0.4166898, pem=-0.25454172,
similarity=17.908804]
35: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.61690456,
final_local_scores=0.5913043, pem=0.0, similarity=21.735413]
36: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.5898114,
final_local_scores=0.5617091, pem=0.0, similarity=20.852234]
37: 6838977 Dhahran
    [pem_pos=0, final_scores=0.3314739,
final_local_scores=0.32569772, pem=0.0, similarity=13.809182]
38: 51554 Riyadh
    [pem_pos=0, final_scores=0.37984723,
final_local_scores=0.35548982, pem=0.0, similarity=14.698238]
39: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.41987664,
final_local_scores=0.37608117, pem=-0.25454172,
similarity=16.696962]
40: 367801 Arlen_Specter
    [pem_pos=0, final_scores=0.39961427,
final_local_scores=0.4325847, pem=0.0, similarity=16.998903]
41: 3434750 United_States
    [pem_pos=0, final_scores=0.50170237,
final_local_scores=0.4922377, pem=-0.073457584,
similarity=19.178541]
```

```
42: 3434750 United_States
    [pem_pos=0, final_scores=0.49115676,
final_local_scores=0.4807182, pem=-0.14075686,
similarity=19.20076]
43: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.46343964,
final_local_scores=0.42366713, pem=-0.25454172,
similarity=18.117022]

Sample: b'475_Hijacked'
Hijacked [1 1 Sudanese 1] plane leaves [2 2 Cyprus 2] for [3
3 Britain 3] . [4 4 LARNACA 4] , [5 5 Cyprus 5] 1996-08-27 A [
6 6 Sudan Airways 6] plane with 199 passengers and crew which
was hijacked to [7 7 Cyprus 7] took off from [8 8 Larnaca 8]
airport after refuelling and headed to [9 9 Britain 9] early
on Tuesday , witnesses said . Police said an unknown number
of hijackers were on board the Airbus 310 which was hijacked
on its way from [10 10 Khartoum 10] to [11 11 Amman 11] in [12
12 Jordan 12] . One had threatened to blow it up unless it
was refuelled and they were taken to [13 13 London 13] where
they intended to surrender and seek political asylum .
```

```
1: 27421 Sudan
2: 5593 Cyprus
3: 31717 United_Kingdom
4: 287852 Larnaca
5: 5593 Cyprus
6: 383994 Sudan_Airways
7: 5593 Cyprus
8: 287852 Larnaca
9: 31717 United_Kingdom
10: 17320 Khartoum
11: 45007 Amman
12: 7515964 Jordan
13: 17867 London
1: 27421 Sudan
    [pem_pos=0, final_scores=0.51592475,
final_local_scores=0.5016336, pem=-0.19893567,
similarity=20.141302]
2: 5593 Cyprus
    [pem_pos=0, final_scores=0.59937793,
final_local_scores=0.5618732, pem=-0.3752463,
similarity=22.897778]
3: 31717 United_Kingdom
    [pem_pos=1, final_scores=0.5232658,
final_local_scores=0.51016814, pem=-0.61114657,
similarity=22.637657]
4: 287852 Larnaca
    [pem_pos=0, final_scores=0.4445861,
final_local_scores=0.4477077, pem=0.0, similarity=17.450205]
5: 5593 Cyprus
    [pem_pos=0, final_scores=0.5184857,
final_local_scores=0.47351074, pem=-0.3752463,
```

```
similarity=20.260864]
6: 383994 Sudan_Airways
    [pem_pos=0, final_scores=0.5041831,
final_local_scores=0.6031136, pem=0.0, similarity=22.087826]
7: 5593 Cyprus
    [pem_pos=0, final_scores=0.65899736,
final_local_scores=0.62699836, pem=-0.3752463,
similarity=24.841242]
8: 287852 Larnaca
    [pem_pos=0, final_scores=0.4056037,
final_local_scores=0.40512538, pem=0.0, similarity=16.179462]
9: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.49216324,
final_local_scores=0.47619337, pem=-0.61114657,
similarity=21.623783]
10: 17320 Khartoum
    [pem_pos=0, final_scores=0.41607326,
final_local_scores=0.40270287, pem=-0.03966028,
similarity=16.32285]
11: 45007 Amman
    [pem_pos=0, final_scores=0.47353923,
final_local_scores=0.4565103, pem=0.0, similarity=17.71289]
12: 7515964 Jordan
    [pem_pos=0, final_scores=0.4497462,
final_local_scores=0.4102198, pem=-0.53647494,
similarity=19.248924]
13: 17867 London
    [pem_pos=0, final_scores=0.48559433,
final_local_scores=0.4815083, pem=-0.09086737,
similarity=18.953032]
```

Sample: b'476_Kurds'

[1 1 Kurds 1] clash again in [2 2 Iraq 2] , dozens reported dead . [3 3 ANKARA 3] 1996-08-26 Heavy fighting broke out between two rival [4 4 Kurdish 4] factions in northern [5 5 Iraq 5] at midnight Sunday and at least 29 people were killed , one of the groups said on Monday . The [6 6 Kurdistan Democratic Party 6] ([7 7 KDP 7]) said the [8 8 Patriotic Union of Kurdistan 8] ([9 9 PUK 9]) had broken the U.S.-brokered ceasefire agreed between the two parties last week . The factions reached a ceasefire on Friday after a week of fierce fighting which had put an end to a truce agreed a year earlier . " The [10 10 PUK 10] leadership who pledged to end fighting and cooperate with the latest [11 11 U.S. 11] initiative started a major military offensive against [12 12 KDP 12] positions , " the [13 13 KDP 13] said in a statement . It said heavy fighting started at midnight in the region dividing the two warring factions , with the [14 14 PUK 14] aiming to break through to [15 15 KDP 15] 's headquarters in Salahuddin . The [16 16 KDP 16] said 29 [17 17 PUK 17] fighters were killed in the attack . It did not provide details of [18 18 KDP 18] casualties and a [19 19 PUK 19]

spokesman was not immediately available for comment . The statement said the [20 20 PUK 20] resumed its attack on Monday morning on [21 21 KDP 21] positions near [22 22 Rawand ouz 22] and indiscriminately shelled the town of Dayana , killing a priest and injuring some civilians . The fighting has threatened a U.S.-led peace plan to unite the mountainous [1 1 Kurdish 1] region in northern [23 23 Iraq 23] against President [24 24 Saddam Hussein 24] . A U.S.-led air force has protected [2 Iraqi Kurds 2] against attack from [25 25 Ba ghdad 25] since shortly after the [26 26 Gulf War 26] in 1991 .

1: 17068 Kurdish_people
2: 7515928 Iraq
3: 802 Ankara
4: 17068 Kurdish_people
5: 7515928 Iraq
6: 25659771 Kurdistan_Democratic_Party
7: 25659771 Kurdistan_Democratic_Party
8: 70789 Patriotic_Union_of_Kurdistan
9: 70789 Patriotic_Union_of_Kurdistan
10: 70789 Patriotic_Union_of_Kurdistan
11: 3434750 United_States
12: 25659771 Kurdistan_Democratic_Party
13: 25659771 Kurdistan_Democratic_Party
14: 70789 Patriotic_Union_of_Kurdistan
15: 25659771 Kurdistan_Democratic_Party
16: 25659771 Kurdistan_Democratic_Party
17: 70789 Patriotic_Union_of_Kurdistan
18: 25659771 Kurdistan_Democratic_Party
19: 70789 Patriotic_Union_of_Kurdistan
20: 70789 Patriotic_Union_of_Kurdistan
21: 25659771 Kurdistan_Democratic_Party
22: 4063907 Rowanduz
23: 7515928 Iraq
24: 29490 Saddam_Hussein
25: 4492 Baghdad
26: 182000 Gulf_War
1: 17068 Kurdish_people
[winner: 80777 Kurdistan, pem_pos=9,
final_scores=0.21876727, final_local_scores=0.18081486,
pem=-2.8001814, similarity=24.713411]
gt: 17068 Kurdish_people, pem_pos=1,
final_scores=0.051513925, final_local_scores=-0.013311446,
pem=-0.713598, similarity=7.57313]
1: 17068 Kurdish_people
[pem_pos=0, final_scores=0.25795412,
final_local_scores=0.21451989, pem=0.0, similarity=10.491414]
2: 7515928 Iraq
[pem_pos=0, final_scores=0.5875454,
final_local_scores=0.5716646, pem=-0.25379938,
similarity=22.529524]
3: 802 Ankara

```
[pem_pos=0, final_scores=0.3551975,
final_local_scores=0.38733, pem=0.0, similarity=15.648412]
4: 17068 Kurdish_people
    [pem_pos=1, final_scores=0.109476134,
final_local_scores=0.052330285, pem=-0.713598,
similarity=9.53201]
5: 7515928 Iraq
    [pem_pos=0, final_scores=0.60731536,
final_local_scores=0.5932602, pem=-0.25379938,
similarity=23.173983]
6: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.653391,
final_local_scores=0.6154391, pem=0.0, similarity=22.455645]
7: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.6154472,
final_local_scores=0.5739913, pem=0.0, similarity=21.218758]
8: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.64044285,
final_local_scores=0.61217934, pem=0.0, similarity=22.358366]
9: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.5975226,
final_local_scores=0.5652955, pem=0.0, similarity=20.959259]
10: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.58298683,
final_local_scores=0.5494174, pem=0.0, similarity=20.485424]
11: 3434750 United_States
    [pem_pos=0, final_scores=0.36344796,
final_local_scores=0.4337222, pem=-0.14075686,
similarity=17.798307]
12: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.5659438,
final_local_scores=0.5199163, pem=0.0, similarity=19.605053]
13: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.49468416,
final_local_scores=0.44207603, pem=0.0, similarity=17.282143]
14: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.6065247,
final_local_scores=0.5751289, pem=0.0, similarity=21.252707]
15: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.7052931,
final_local_scores=0.6721343, pem=0.0, similarity=24.147541]
16: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.50479275,
final_local_scores=0.4531181, pem=0.0, similarity=17.61166]
17: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.56587404,
final_local_scores=0.5307242, pem=0.0, similarity=19.927582]
18: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.57944614,
final_local_scores=0.5346655, pem=0.0, similarity=20.045198]
19: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.54128176,
```

```
final_local_scores=0.5038609, pem=0.0, similarity=19.125927]
20: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.6225555,
final_local_scores=0.59264016, pem=0.0, similarity=21.775278]
21: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.6048597,
final_local_scores=0.56242603, pem=0.0, similarity=20.873627]
22: 4063907 Rowanduz
    [pem_pos=0, final_scores=0.36443365,
final_local_scores=0.43893284, pem=0.0, similarity=17.188345]
23: 7515928 Iraq
    [pem_pos=0, final_scores=0.5734218,
final_local_scores=0.5562366, pem=-0.25379938,
similarity=22.069124]
24: 29490 Saddam_Hussein
    [pem_pos=0, final_scores=0.4997353,
final_local_scores=0.482678, pem=0.0, similarity=18.493788]
25: 4492 Baghdad
    [pem_pos=0, final_scores=0.4112847,
final_local_scores=0.38819665, pem=0.0, similarity=15.674274]
26: 182000 Gulf_War
    [pem_pos=0, final_scores=0.29200166,
final_local_scores=0.29998013, pem=0.0, similarity=13.0417185]
1: 80777 Kurdistan
    [pem_pos=9, final_scores=0.21876727,
final_local_scores=0.18081486, pem=-2.8001814,
similarity=24.713411]
2: 679693 Iraqi_Kurdistan
    [pem_pos=0, final_scores=0.23734353,
final_local_scores=0.17773792, pem=0.0, similarity=9.393767]
```

Sample: b'477_Egypt'

[1 1 Egypt 1] to press [2 2 Britain 2] over [3 3 Islamist 3] conference . [4 4 CAIRO 4] 1996-08-26 [5 5 Egypt 5] will tell [6 6 Britain 6] it is concerned about a meeting of Islamists to be held in [7 7 London 7] soon , Foreign Minister [8 8 Amr Moussa 8] said on Monday . " There is a question mark over this issue . We , and many other countries , do n't understand this ([9 9 Britain 9] 's) position , " [10 10 Mo ussa 10] told reporters . " [11 11 Egypt 11] will contact the [12 12 British 12] government to find out the truth of the matter and to discuss the possible consequences of such an unfortunate step , " he added . [13 13 Egyptian 13] government newspapers have criticised [14 14 Britain 14] for allowing Islamists , whom they brand as " terrorists " , to hold their conference , saying the meeting will be a chance for dangerous [15 15 Moslem 15] militants to plot against their countries of origin . It is not clear when the conference will be held . About 1,000 people have been killed in [16 16 Egypt 16] since [17 17 Islamic 17] militants took up arms in 1992 in an attempt to overthrow the government and set up a strict [18 18 Islamic 18] state . [19 19 Cairo 19]

says several [20 20 Egyptian 20] militants on the run from death sentences or convictions for violent attacks at home have taken shelter in [21 21 Britain 21] .

1: 8087628 Egypt
2: 31717 United_Kingdom
3: 6037917 Islam
4: 6293 Cairo
5: 8087628 Egypt
6: 31717 United_Kingdom
7: 17867 London
8: 201993 Amr_Moussa
9: 31717 United_Kingdom
10: 201993 Amr_Moussa
11: 8087628 Egypt
12: 31717 United_Kingdom
13: 8087628 Egypt
14: 31717 United_Kingdom
15: 6037917 Islam
16: 8087628 Egypt
17: 6037917 Islam
18: 6037917 Islam
19: 6293 Cairo
20: 8087628 Egypt
21: 31717 United_Kingdom
1: 8087628 Egypt
[pem_pos=0, final_scores=0.55602837,
final_local_scores=0.50673926, pem=-0.27480945,
similarity=20.70628]
2: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5891611,
final_local_scores=0.57224756, pem=-0.61114657,
similarity=24.490232]
3: 6037917 Islam
[pem_pos=1, final_scores=0.28439003,
final_local_scores=0.26665327, pem=-1.3455799,
similarity=19.36465]
4: 6293 Cairo
[pem_pos=0, final_scores=0.5817767,
final_local_scores=0.55701965, pem=-0.15211283,
similarity=21.539505]
5: 8087628 Egypt
[pem_pos=0, final_scores=0.5858752,
final_local_scores=0.53934246, pem=-0.27480945,
similarity=21.679224]
6: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5614361,
final_local_scores=0.5419622, pem=-0.61114657,
similarity=23.586456]
7: 17867 London
[pem_pos=0, final_scores=0.47421962,
final_local_scores=0.47300786, pem=-0.09086737,
similarity=18.699362]

```
8: 201993 Amr_Moussa
    [pem_pos=0, final_scores=0.578995,
final_local_scores=0.63635224, pem=0.0, similarity=23.079733]
9: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.55241495,
final_local_scores=0.5321079, pem=-0.61114657,
similarity=23.292383]
10: 201993 Amr_Moussa
    [pem_pos=0, final_scores=0.62724066,
final_local_scores=0.7014007, pem=0.0, similarity=25.020908]
11: 8087628 Egypt
    [pem_pos=0, final_scores=0.5493786,
final_local_scores=0.49947542, pem=-0.27480945,
similarity=20.489513]
12: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.57234496,
final_local_scores=0.55387837, pem=-0.48936236,
similarity=23.279778]
13: 8087628 Egypt
    [pem_pos=0, final_scores=0.51674825,
final_local_scores=0.46383172, pem=-1.052751,
similarity=23.6564]
14: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5626977,
final_local_scores=0.54334027, pem=-0.61114657,
similarity=23.62758]
15: 6037917 Islam
    [pem_pos=1, final_scores=0.46139294,
final_local_scores=0.460002, pem=-1.2150909,
similarity=24.424942]
16: 8087628 Egypt
    [pem_pos=0, final_scores=0.5249062,
final_local_scores=0.47274303, pem=-0.27480945,
similarity=19.691765]
17: 6037917 Islam
    [pem_pos=0, final_scores=0.39062375,
final_local_scores=0.38269746, pem=-0.21141872,
similarity=16.659897]
18: 6037917 Islam
    [pem_pos=0, final_scores=0.4010784,
final_local_scores=0.39411753, pem=-0.21141872,
similarity=17.000694]
19: 6293 Cairo
    [pem_pos=0, final_scores=0.40505928,
final_local_scores=0.3639828, pem=-0.15211283,
similarity=15.7789]
20: 8087628 Egypt
    [pem_pos=0, final_scores=0.4681056,
final_local_scores=0.41069698, pem=-1.052751,
similarity=22.070753]
21: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5154536,
```

```
final_local_scores=0.49173325, pem=-0.61114657,  
similarity=22.087524]
```

Sample: b'478_Israeli'

[1 1 Israeli 1] army ransacks Bedouin Palestinian camp . AL-MUNTAR , West Bank 1996-08-26 [2 2 Israeli 2] security forces ransacked a [3 3 Bedouin 3] encampment in the West Bank on Monday to expel them from an area which [4 4 Palestinians 4] say had been earmarked for [5 5 Jewish 5] settlement expansion , residents said . They said soldiers stole a gold necklace and about \$ 2,000 in [6 6 Israeli 6] currency from an elderly woman and her daughter-in-law while rummaging through their luggage before destroying family shacks and animal barns . " They rammed our shacks with jeeps and destroyed the shack over my baby , " said 25-year-old Amina Muhammad . " He was saved only by a miracle . " A spokesman for [7 7 Israel 7] 's civil administration in the West Bank said the [8 8 Bedouins 8] were moved because they were encamped on an [9 9 Israeli 9] army firing zone . The spokesman Peter Lerner said he knew nothing about soldiers having stolen anything from the [10 10 Palestinians 10] . [11 11 Israeli 11] security forces have been pursuing Bedouin Palestinians living in the desolate wilderness between [12 12 East Jerusalem 12] and the [13 13 Dead Sea 13] , where several [14 14 Jewish 14] settlements have been established . The [15 15 Israeli 15] army also uses the area for military training .

1: 9282173 Israel

2: 9282173 Israel

3: 61770 Bedouin

4: 1114732 Palestine

5: 25955086 Jews

6: 9282173 Israel

7: 9282173 Israel

8: 61770 Bedouin

9: 9282173 Israel

10: 1114732 Palestine

11: 9282173 Israel

12: 343508 East_Jerusalem

13: 8530 Dead_Sea

14: 25955086 Jews

15: 9282173 Israel

1: 9282173 Israel

```
[pem_pos=0, final_scores=0.59643227,  
final_local_scores=0.5391565, pem=-0.34790626,  
similarity=22.071186]
```

2: 9282173 Israel

```
[pem_pos=0, final_scores=0.5459326,  
final_local_scores=0.48399335, pem=-0.34790626,  
similarity=20.425009]
```

3: 61770 Bedouin

```
[pem_pos=0, final_scores=0.39321977,
```

```
final_local_scores=0.39665192, pem=0.0, similarity=15.926597]
4: 1114732 Palestine
    [pem_pos=1, final_scores=0.36712855,
final_local_scores=0.30681926, pem=-1.6411982,
similarity=22.170902]
5: 25955086 Jews
    [pem_pos=1, final_scores=0.44739187,
final_local_scores=0.4243415, pem=-0.80476934,
similarity=21.12937]
6: 9282173 Israel
    [pem_pos=0, final_scores=0.5224762,
final_local_scores=0.45837075, pem=-0.34790626,
similarity=19.660378]
7: 9282173 Israel
    [pem_pos=0, final_scores=0.5793707,
final_local_scores=0.5205193, pem=-0.21279564,
similarity=20.780262]
8: 61770 Bedouin
    [pem_pos=0, final_scores=0.36406952,
final_local_scores=0.36480975, pem=0.0, similarity=14.976364]
9: 9282173 Israel
    [pem_pos=0, final_scores=0.5587455,
final_local_scores=0.49798948, pem=-0.34790626,
similarity=20.842682]
10: 1114732 Palestine
    [pem_pos=1, final_scores=0.36865968,
final_local_scores=0.3084917, pem=-1.6411982,
similarity=22.220812]
11: 9282173 Israel
    [pem_pos=0, final_scores=0.5496529,
final_local_scores=0.4880572, pem=-0.34790626,
similarity=20.54628]
12: 343508 East_Jerusalem
    [pem_pos=0, final_scores=0.33110017,
final_local_scores=0.31783196, pem=0.0, similarity=13.574452]
13: 8530 Dead_Sea
    [pem_pos=0, final_scores=0.39092547,
final_local_scores=0.42219877, pem=0.0, similarity=16.688967]
14: 25955086 Jews
    [pem_pos=1, final_scores=0.45950633,
final_local_scores=0.43757468, pem=-0.80476934,
similarity=21.524275]
15: 9282173 Israel
    [pem_pos=0, final_scores=0.6158361,
final_local_scores=0.56035227, pem=-0.34790626,
similarity=22.70371]
```

Sample: b'479_U.S.'

[1 1 U.S. 1] base metals and scrap prices - August 26 . [2 2
U.S. 2] premiums added to LME official cash settlement
price , except for copper , which is added to the [3 3 COMEX 3
] spot month . Premium includes price for delivery to

consumer 's works . -- Aluminum : [4 4 Western 4] grade
3.25-3.75 cents / pound [5 5 Russian 5] grade A7E 3.25-3.75
cents / pound [6 6 Russian 6] grade A0 nominal 2.00-2.25
cents / pound Zinc : Special high grade SHG 5.50-6.00 cents /
pound Lead 3.50-4.00 cents / pound Tin (Grade A) 6.5-8.5
cents / pound (low lead , 50 ppm) 9.0-10.5 cents / pound
Nickel (melting grade) 9.0-12.0 cents / pound Copper (high
grade cathode) 2.50-3.0 cents / pound -- Aluminum alloy
(A380 grade) , [7 7 Midwest 7] and East coast , delivered
to consumer 65-66 cents / pound -- Aluminum scrap , [8 8 Midw
est 8] and East coast average price , delivered to consumer
Old Sheet and Cast metal 42 to 44 cents / pound Turnings ,
clean and dry 43 to 44 cents / pound Mixed low-copper clips
48 to 49 cents / pound -- Copper scrap , [9 9 Midwest 9] and
East coast average price , delivered to consumer No2 Refined
75 to 77 cents / pound Nol Bare Bright 91 to 92 cents / pound
Nol Burnt 87 to 90 cents / pound -- Lead batteries ,
delivered consumer 4.5 to 6.0 cents / pound [10 10 U.S. 10]
Producer list / transaction prices -- [11 11 Alcan 11]
aluminum , [12 12 U.S. Midwest 12] (effective date : August
1 , 1996) P1020 ingot 75 cents / pound extrusion billet 85
cents / pound [13 13 Noranda 13] aluminum , [14 14 U.S. Midwe
st 14] (effective date:August 1 , 1996) ingot 75 cents /
pound extrusion billet 85 cents / pound RSR pure lead price
(effective date : March 20 , 1996) 52 cents / pound Doe Run
pure lead price (effective date : August 14 , 1996) 50
cents / pound [15 15 ASARCO 15] pure lead (effective date :
August 1 , 1996) premium over LME cash 7.5 cents / pound --
([16 16 New York 16] commodities desk 212 859 1646)
1: 3434750 United_States
2: 3434750 United_States
3: 469729 New_York_Mercantile_Exchange
4: 21208200 Western_world
5: 25391 Russia
6: 25391 Russia
7: 104697 Midwestern_United_States
8: 104697 Midwestern_United_States
9: 104697 Midwestern_United_States
10: 3434750 United_States
11: 323202 Rio_Tinto_Alcan
12: 104697 Midwestern_United_States
13: 978322 Noranda_(mining_company)
14: 104697 Midwestern_United_States
15: 541026 Asarco
16: 645042 New_York_City
1: 3434750 United_States
[pem_pos=0, final_scores=0.547262,
final_local_scores=0.514219, pem=-0.14075686,
similarity=20.20049]
2: 3434750 United_States
[pem_pos=0, final_scores=0.4934817,
final_local_scores=0.45547217, pem=-0.14075686,

```
similarity=18.447369]
3: 469729 New_York_Mercantile_Exchange
    [pem_pos=0, final_scores=0.43271893,
final_local_scores=0.5276256, pem=0.0, similarity=19.835114]
4: 21208200 Western_world
    [pem_pos=0, final_scores=0.35613978,
final_local_scores=0.3950321, pem=-1.5917408,
similarity=24.534393]
5: 25391 Russia
    [pem_pos=1, final_scores=0.6853511,
final_local_scores=0.7326883, pem=-1.0515022,
similarity=31.672827]
6: 25391 Russia
    [pem_pos=1, final_scores=0.77221674,
final_local_scores=0.8275758, pem=-1.0515022,
similarity=34.50446]
7: 104697 Midwestern_United_States
    [pem_pos=0, final_scores=0.4836921,
final_local_scores=0.44150895, pem=-0.031208755,
similarity=17.434938]
8: 104697 Midwestern_United_States
    [pem_pos=0, final_scores=0.46684682,
final_local_scores=0.42310804, pem=-0.031208755,
similarity=16.885818]
9: 104697 Midwestern_United_States
    [pem_pos=0, final_scores=0.46177554,
final_local_scores=0.41756845, pem=-0.031208755,
similarity=16.720507]
10: 3434750 United_States
    [pem_pos=0, final_scores=0.45899397,
final_local_scores=0.41779953, pem=-0.14075686,
similarity=17.323141]
11: 323202 Rio_Tinto_Alcan
    [pem_pos=0, final_scores=0.32980984,
final_local_scores=0.41697192, pem=0.0, similarity=16.532988]
12: 104697 Midwestern_United_States
    [pem_pos=0, final_scores=0.44726032,
final_local_scores=0.40171272, pem=0.0, similarity=16.077621]
13: 978322 Noranda_(mining_company)
    [pem_pos=0, final_scores=0.45988017,
final_local_scores=0.5574248, pem=-0.04634327,
similarity=20.976402]
14: 104697 Midwestern_United_States
    [pem_pos=0, final_scores=0.45116895,
final_local_scores=0.40598232, pem=0.0, similarity=16.205034]
15: 541026 Asarco
    [pem_pos=0, final_scores=0.49310338,
final_local_scores=0.58303386, pem=0.0, similarity=21.488605]
16: 645042 New_York_City
    [pem_pos=1, final_scores=0.38582635,
final_local_scores=0.37027723, pem=-1.7906687,
similarity=24.877459]
```

Sample: b'480_Florida'

Florida boy kills himself before starting new school . [1 1 F
ORT LAUDERDALE 1] , Fla. 1996-08-26 A 12-year-old Florida boy hanged himself in his backyard just hours before he was due to start at a new school on Monday , police said . Samuel Graham , who told his family earlier that he was nervous about starting at a new school because he feared teasing about his weight problem , had been due to spend his first day at Parkway Middle School Monday , police said . The boy was last seen alive Sunday night when he joined his two younger brothers and father in a bedtime prayer . Two younger brothers found him hanging from a tree early Monday morning . His father cut him down and tried to revive him but paramedics pronounced him dead when they arrived . The Broward County Sheriff 's Office found a step stool and a flashlight under the tree where the boy was hanged . They said there was no sign of foul play and that investigators believed the death was a suicide .

1: 109028 Fort_Lauderdale,_Florida

1: 109028 Fort_Lauderdale,_Florida

[pem_pos=0, final_scores=0.20334823,
final_local_scores=0.49471426, pem=0.0, similarity=18.852974]

Sample: b'481_CBOT'

CBOT rice closes higher on technical bounce . [1 1 CHICAGO 1]
1996-08-26 CBOT rice futures closed higher on a technical bounce tied to signs the market was oversold , traders said . " It rallied on ideas the market was oversold , " a trader said . Traders said the [2 2 U.S. 2] cash market remains well above the benchmark [3 3 Thai 3] price which limited gains . Rice futures volume was estimated at 450 contracts , up from 202 Friday . Rice options volume was estimated at 50 contracts , down from 56 Friday . Rice futures closed 13 to 16 cents per cwt higher , with September up 16 at \$ 10.28 a cwt . Sam Nelson 312-408-8721

1: 6886 Chicago

2: 3434750 United_States

3: 30128 Thailand

1: 6886 Chicago

[pem_pos=0, final_scores=0.52302855,
final_local_scores=0.54742366, pem=-0.20658566,
similarity=21.549372]

2: 3434750 United_States

[pem_pos=0, final_scores=0.5219657,
final_local_scores=0.49498135, pem=-0.14075686,
similarity=19.626402]

3: 30128 Thailand

[pem_pos=2, final_scores=0.6171663,
final_local_scores=0.6813703, pem=-1.7504287,
similarity=33.942265]

Sample: b'482_Talbott'

[1 1 Strobe_Talbott 1] to meet [2 2 Russian 2] , [3 3 Canadian 3] counterparts . [4 4 WASHINGTON 4] 1996-08-26 Deputy Secretary of State [5 5 Strobe Talbott 5] flew to [6 6 Ottawa 6] on Monday to meet his [7 7 Russian 7] counterpart and discuss a range of bilateral and [8 8 European 8] security issues , the [9 9 State Department 9] said . [10 10 Strobe_Talbott 10] , the second-ranking [11 11 State Department 11] official , was to meet [12 12 Russian 12] Deputy Foreign Minister [13 13 Georgy Mamedov 13] as part of a regular pattern of consultations , the department said . " There is much to discuss on the fall calendar , " acting chief spokesman Glyn Davies said . " There 's a fairly intensive diplomatic calendar coming up in the fall . " He said [14 14 Strobe_Talbott 14] , who was scheduled to return on Tuesday , would also to meet his [15 15 Canadian 15] counterpart , Gordon Smith , in [16 16 Ottawa 16] for talks that would include the situation in [17 17 Haiti 17] .

1: 917589 Strobe_Talbott
2: 25391 Russia
3: 5042916 Canada
4: 108956 Washington,_D.C.
5: 917589 Strobe_Talbott
6: 22219 Ottawa
7: 25391 Russia
8: 9239 Europe
9: 31975 United_States_Department_of_State
10: 917589 Strobe_Talbott
11: 31975 United_States_Department_of_State
12: 25391 Russia
13: 3037636 Georgiy_Mamedov
14: 917589 Strobe_Talbott
15: 5042916 Canada
16: 22219 Ottawa
17: 13373 Haiti
1: 917589 Strobe_Talbott
[pem_pos=0, final_scores=1.079539,
final_local_scores=1.1483285, pem=0.0, similarity=38.358124]
2: 25391 Russia
[pem_pos=1, final_scores=0.600146,
final_local_scores=0.6356916, pem=-1.0515022,
similarity=28.77825]
3: 5042916 Canada
[pem_pos=0, final_scores=0.5508792,
final_local_scores=0.57646435, pem=-0.41303828,
similarity=23.538725]
4: 108956 Washington,_D.C.
[pem_pos=0, final_scores=0.5849376,
final_local_scores=0.59592, pem=-1.7496481,
similarity=31.388012]
5: 917589 Strobe_Talbott
[pem_pos=0, final_scores=0.75049216,
final_local_scores=0.77762973, pem=0.0, similarity=27.295734]

```
6: 22219 Ottawa
    [pem_pos=0, final_scores=0.49569887,
final_local_scores=0.5158183, pem=-0.34965137,
similarity=21.384218]
7: 25391 Russia
    [pem_pos=1, final_scores=0.6504484,
final_local_scores=0.6906394, pem=-1.0515022,
similarity=30.418003]
8: 9239 Europe
    [pem_pos=0, final_scores=0.38002145,
final_local_scores=0.41732252, pem=-0.25557294,
similarity=17.933294]
9: 31975 United_States_Department_of_State
    [pem_pos=0, final_scores=0.51393515,
final_local_scores=0.5111958, pem=0.0, similarity=19.344814]
10: 917589 Strobe_Talbott
    [pem_pos=0, final_scores=1.035325,
final_local_scores=1.1000316, pem=0.0, similarity=36.91685]
11: 31975 United_States_Department_of_State
    [pem_pos=0, final_scores=0.39752573,
final_local_scores=0.38403618, pem=0.0, similarity=15.550119]
12: 25391 Russia
    [pem_pos=1, final_scores=0.8507282,
final_local_scores=0.9094146, pem=-1.0515022,
similarity=36.946693]
13: 3037636 Georgiy_Mamedov
    [pem_pos=0, final_scores=0.5207504,
final_local_scores=0.602069, pem=0.0, similarity=22.056654]
14: 917589 Strobe_Talbott
    [pem_pos=0, final_scores=0.9854513,
final_local_scores=1.045552, pem=0.0, similarity=35.29107]
15: 5042916 Canada
    [pem_pos=0, final_scores=0.45988292,
final_local_scores=0.4770648, pem=-0.41303828,
similarity=20.572443]
16: 22219 Ottawa
    [pem_pos=0, final_scores=0.50963885,
final_local_scores=0.5310457, pem=-0.34965137,
similarity=21.838634]
17: 13373 Haiti
    [pem_pos=0, final_scores=0.3955928,
final_local_scores=0.4270543, pem=-0.26845667,
similarity=18.293774]
```

Sample: b'483_U.S.'

```
[1 1 U.S. Midwest 1] hog market seen steady Tuesday -
trade . [2 2 CHICAGO 2] 1996-08-26 [3 3 Midwest 3] direct
cash hog prices Tuesday were seen steady following strong
demand Monday that lifted prices as much as $ 1.50 per cwt in
some areas , livestock dealers said . The demand was sparked
by Saturday 's active slaughter , which left some packers
short on supplies to get Monday operations started , they
```

said . Top prices in Iowa / southern [4 4 Minnesota 4] Tuesday were expected to range from mostly \$ 59.50 to \$ 60.00 , steady following a \$ 0.50 - \$ 1.50 jump Monday . [5 5 Illinois 5] tops were seen matching Monday 's at \$ 59.00 with tops in [6 6 Indiana 6] at \$ 57.50 . However , [7 7 USDA 7] reported tops of \$ 58.00 in [8 8 Illinois 8] and in Iowa / southern [9 9 Minnesota 9] , \$ 60.00 on some hogs Monday . Attempts to move prices higher again Tuesday could be offset by expected increased hog marketings this week , sources said . Producers were expected to ship as many hogs as they can ahead of the [10 10 Labor Day 10] holiday weekend , they said . Demand for hogs is expected to be light by the end of the week , as the industry prepares for [11 11 Labor Day 11] , Monday , September 1 when most [12 12 U.S. 12] packers will be closed . -- Bob Janis 312-983-7347--

1: 104697 Midwestern_United_States
2: 6886 Chicago
3: 104697 Midwestern_United_States
4: 19590 Minnesota
5: 14849 Illinois
6: 21883857 Indiana
7: 70896 United_States_Department_of_Agriculture
8: 14849 Illinois
9: 19590 Minnesota
10: 19156272 Labor_Day
11: 19156272 Labor_Day
12: 3434750 United_States
1: 104697 Midwestern_United_States
[pem_pos=0, final_scores=0.46673685,
final_local_scores=0.4220702, pem=0.0, similarity=16.68513]
2: 6886 Chicago
[pem_pos=0, final_scores=0.5937045,
final_local_scores=0.5647872, pem=-0.20658566,
similarity=22.067535]
3: 104697 Midwestern_United_States
[pem_pos=0, final_scores=0.4771979,
final_local_scores=0.4334973, pem=-0.031208755,
similarity=17.195856]
4: 19590 Minnesota
[pem_pos=0, final_scores=0.5193828,
final_local_scores=0.48050302, pem=-0.21388707,
similarity=19.592031]
5: 14849 Illinois
[pem_pos=0, final_scores=0.558216,
final_local_scores=0.5067544, pem=-0.16995399,
similarity=20.13651]
6: 21883857 Indiana
[pem_pos=0, final_scores=0.496494,
final_local_scores=0.45667303, pem=-0.27230176,
similarity=19.198566]
7: 70896 United_States_Department_of_Agriculture
[pem_pos=0, final_scores=0.5304187,

```
final_local_scores=0.6015563, pem=0.0, similarity=22.041353]
8: 14849 Illinois
    [pem_pos=0, final_scores=0.51377654,
final_local_scores=0.45821106, pem=-0.16995399,
similarity=18.687881]
9: 19590 Minnesota
    [pem_pos=0, final_scores=0.52996427,
final_local_scores=0.49206167, pem=-0.21388707,
similarity=19.936966]
10: 19156272 Labor_Day
    [pem_pos=0, final_scores=0.40020388,
final_local_scores=0.40456688, pem=0.0, similarity=16.162796]
11: 19156272 Labor_Day
    [pem_pos=0, final_scores=0.48485512,
final_local_scores=0.49703544, pem=0.0, similarity=18.922241]
12: 3434750 United_States
    [pem_pos=0, final_scores=0.5431211,
final_local_scores=0.5240229, pem=-0.14075686,
similarity=20.493057]
```

Sample: b'484_Gore'

[1 1 Gore 1] presents new image of lead attack dog . Michael Posner [2 2 CHICAGO 2] 1996-08-26 No more mild-mannered and meek [3 3 Al Gore 3] , the vice president and likely heir apparent to President [4 4 Bill Clinton 4] , emerged on Monday as the new [5 5 Democratic 5] attack dog leading a front-line assault on [6 6 Bob Dole 6] and House Speaker [7 7 Newt Gingrich 7] . While [8 8 Clinton 8] takes a train trip to the [9 9 Democratic Convention 9] that will renominate him on Wednesday for a second term , [10 10 Gore 10] is the bright star in the convention city now . Crowds anxiously wait for his appearances , thrusting out hands for him to grip as they scream " 12 more years " -- a wishful hope to eight years of [11 11 Gore 11] after a [12 12 Clinton 12] re-election . Over a 15-hour span from Sunday evening to Monday morning , so many people jammed [13 13 Gore 13] events that the fire marshal stopped members of [14 14 Congress 14] , reporters and others from entering a pro-Israel rally and a meeting of the [15 15 New York 15] delegation . In his appearances , the often stiff and wooden [16 16 Gore 16] seemed transformed into a new energetic , gesturing " pol " as he ripped into [17 17 Republican 17] presidential nominee [18 18 Dole 18] and [19 19 Gingrich 19] , who has emerged as the favorite right-wing foil of [20 20 Democrats 20] . [21 21 Gore 21] told a roaring labor rally of about 1,000 union workers that [22 22 Dole 22] and [23 23 Gingrich 23] were the virtual personification of evil , without even mentioning that former Housing Secretary [24 24 Jack Kemp 24] is [25 25 Dole 25] 's running mate . " With equal measures of ignorance and audacity this two-headed monster of [26 26 Dole 26] and [27 27 Gingrich 27] has been launching an all out assault on decades of progress of behalf of working men and women , " [28 28 Gore

28] said to whoops of " 12 more years . " " They want to drive you out of politics and they ca n't , " he added . " They want to silence your voices in elections . " At a downtown hotel Monday , the 48-year-old [29 29 Gore 29] gave [30 30 Wisconsin 30] delegates a taste again of the new [31 31 Gore 31] . " I want you to ask this question . What would [32 32 Wisconsin 32] face if the same extremist coalition , the Gingrich / Dole Congress , also controlled the executive branch ? " [33 33 Gore 33] said . Noting the next presidential term will probably see two or three [34 34 Supreme Court 34] justice nominations , he warned , " Their extremist agenda would come out of the Gingrich Congress , into and through the Dole White House , down through the [35 35 Supreme Court 35] . " He painted a scene of horrors he saw if [36 36 Republicans 36] controlled the [37 37 White House 37] and the [38 38 Congress 38] they hold now . " Our personal and religious liberties would be at risk . [39 39 Medicare 39] would be at risk of withering on the vine . Medicaid would be at risk of being taken away from poor children . Education would be at risk . The environment would be at risk from the polluters that control that coalition , " he said . His new style mixed an assault on [40 40 Republicans 40] with what he sees as " productive opportunity , a vision for the future that lifts up working families , that builds [41 41 Wisconsin 41] stronger , that builds [42 42 America 42] stronger . With your help we can re-elect [43 43 Bill Clinton 43] . " [44 44 Dole 44] 's tax cut plans were called " *deja-voodoo* economics It 's a warmed-over plan that failed and drove our economy into a ditch . We got burned once and we wo n't let that happen to our nation again . " [45 45 Gore 45] ridiculed [46 46 Dole 46] 's defense of the tobacco industry , praised [47 47 Clinton 47] for " courage " in advancing regulations of it . To the delegates of [48 48 Wisconsin 48] , a leading dairy state , [49 49 Gore 49] had the right audience in poking fun at [50 50 Dole 50] 's comparison to addiction to tobacco to those who ca n't tolerate milk . " Some people say milk is bad for you , " [51 51 Gore 51] said as [52 52 Wisconsin 52] delegates held aloft a plastic inflated cow . He praised [53 53 Clinton 53] for vetoing the [54 54 Republican 54] [55 55 Congress 55] ' attempt to repeal parts of the Clean Air Act , and criticized [56 56 Republicans 56] for " slashing money to combat drugs in schools . " " The choice has never been starker , the stakes today have never been higher , " [57 57 Gore 57] said . " We 're not going to let them get away with that . "

- 1: 5042706 Al_Gore
- 2: 6886 Chicago
- 3: 5042706 Al_Gore
- 4: 3356 Bill_Clinton
- 5: 5043544 Democratic_Party_(United_States)
- 6: 55846 Bob_Dole
- 7: 179187 Newt_Gingrich

```
8: 3356 Bill_Clinton
9: 468102 Democratic_National_Convention
10: 5042706 Al_Gore
11: 5042706 Al_Gore
12: 3356 Bill_Clinton
13: 5042706 Al_Gore
14: 31756 United_States_Congress
15: 645042 New_York_City
16: 5042706 Al_Gore
17: 32070 Republican_Party_(United_States)
18: 55846 Bob_Dole
19: 179187 Newt_Gingrich
20: 5043544 Democratic_Party_(United_States)
21: 5042706 Al_Gore
22: 55846 Bob_Dole
23: 179187 Newt_Gingrich
24: 16200 Jack_Kemp
25: 55846 Bob_Dole
26: 55846 Bob_Dole
27: 179187 Newt_Gingrich
28: 5042706 Al_Gore
29: 5042706 Al_Gore
30: 33127 Wisconsin
31: 5042706 Al_Gore
32: 33127 Wisconsin
33: 5042706 Al_Gore
34: 31737 Supreme_Court_of_the_United_States
35: 31737 Supreme_Court_of_the_United_States
36: 32070 Republican_Party_(United_States)
37: 33057 White_House
38: 31756 United_States_Congress
39: 58353 Medicare_(United_States)
40: 32070 Republican_Party_(United_States)
41: 33127 Wisconsin
42: 3434750 United_States
43: 3356 Bill_Clinton
44: 55846 Bob_Dole
45: 5042706 Al_Gore
46: 55846 Bob_Dole
47: 3356 Bill_Clinton
48: 33127 Wisconsin
49: 5042706 Al_Gore
50: 55846 Bob_Dole
51: 5042706 Al_Gore
52: 33127 Wisconsin
53: 3356 Bill_Clinton
54: 32070 Republican_Party_(United_States)
55: 31756 United_States_Congress
56: 32070 Republican_Party_(United_States)
57: 5042706 Al_Gore
1: 5042706 Al_Gore
    [pem_pos=0, final_scores=0.8242555,
```

```
final_local_scores=0.7965397, pem=0.0, similarity=27.860046]
2: 6886 Chicago
    [pem_pos=0, final_scores=0.60183156,
final_local_scores=0.649566, pem=-0.20658566,
similarity=24.597504]
3: 5042706 Al_Gore
    [pem_pos=0, final_scores=0.491131,
final_local_scores=0.43199432, pem=0.0, similarity=16.981285]
4: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.5554793,
final_local_scores=0.50972694, pem=0.0, similarity=19.300982]
5: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.47762048,
final_local_scores=0.4541921, pem=-0.076093495,
similarity=18.05752]
6: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.69196373,
final_local_scores=0.66955686, pem=0.0, similarity=24.070625]
7: 179187 Newt_Gingrich
    [pem_pos=0, final_scores=0.5122987,
final_local_scores=0.49063498, pem=0.0, similarity=18.73124]
8: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.84445304,
final_local_scores=0.8262429, pem=0.0, similarity=28.74645]
9: 468102 Democratic_National_Convention
    [pem_pos=0, final_scores=0.41239983,
final_local_scores=0.39472264, pem=-0.23190276,
similarity=17.130146]
10: 5042706 Al_Gore
    [pem_pos=0, final_scores=0.7666052,
final_local_scores=0.73356557, pem=0.0, similarity=25.980772]
11: 5042706 Al_Gore
    [pem_pos=0, final_scores=0.7996772,
final_local_scores=0.7696915, pem=0.0, similarity=27.058844]
12: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.84831184,
final_local_scores=0.83045805, pem=0.0, similarity=28.872238]
13: 5042706 Al_Gore
    [pem_pos=0, final_scores=0.849655,
final_local_scores=0.82428473, pem=0.0, similarity=28.688013]
14: 31756 United_States_Congress
    [pem_pos=0, final_scores=0.3116824,
final_local_scores=0.29122117, pem=-0.3349529,
similarity=14.601859]
15: 645042 New_York_City
    [pem_pos=1, final_scores=0.46248907,
final_local_scores=0.49250585, pem=-1.7906687,
similarity=28.525003]
16: 5042706 Al_Gore
    [pem_pos=0, final_scores=0.905664,
final_local_scores=0.88546604, pem=0.0, similarity=30.513786]
17: 32070 Republican_Party_(United_States)
```

```
[pem_pos=0, final_scores=0.45775843,
final_local_scores=0.43094534, pem=-0.1310191,
similarity=17.662483]
18: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.8436236,
final_local_scores=0.83635736, pem=0.0, similarity=29.048285]
19: 179187 Newt_Gingrich
    [pem_pos=0, final_scores=0.51132375,
final_local_scores=0.48956996, pem=0.0, similarity=18.699457]
20: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.48884767,
final_local_scores=0.46645606, pem=-0.18007085,
similarity=18.988945]
21: 5042706 Al_Gore
    [pem_pos=0, final_scores=0.740589,
final_local_scores=0.7051468, pem=0.0, similarity=25.1327]
22: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.77852815,
final_local_scores=0.76525044, pem=0.0, similarity=26.926313]
23: 179187 Newt_Gingrich
    [pem_pos=0, final_scores=0.5195053,
final_local_scores=0.49850714, pem=0.0, similarity=18.96616]
24: 16200 Jack_Kemp
    [pem_pos=0, final_scores=0.49106914,
final_local_scores=0.49590576, pem=0.0, similarity=18.88853]
25: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.8463083,
final_local_scores=0.83929, pem=0.0, similarity=29.135801]
26: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.8357052,
final_local_scores=0.82770777, pem=0.0, similarity=28.790163]
27: 179187 Newt_Gingrich
    [pem_pos=0, final_scores=0.4720828,
final_local_scores=0.44670516, pem=0.0, similarity=17.420286]
28: 5042706 Al_Gore
    [pem_pos=0, final_scores=0.8120062,
final_local_scores=0.78315914, pem=0.0, similarity=27.460743]
29: 5042706 Al_Gore
    [pem_pos=0, final_scores=0.7680865,
final_local_scores=0.7351836, pem=0.0, similarity=26.029058]
30: 33127 Wisconsin
    [pem_pos=0, final_scores=0.51459134,
final_local_scores=0.5581966, pem=-0.21268336,
similarity=21.904018]
31: 5042706 Al_Gore
    [pem_pos=0, final_scores=0.7814999,
final_local_scores=0.7498357, pem=0.0, similarity=26.466307]
32: 33127 Wisconsin
    [pem_pos=0, final_scores=0.565058,
final_local_scores=0.6133238, pem=-0.21268336,
similarity=23.549124]
33: 5042706 Al_Gore
```

[pem_pos=0, final_scores=0.8137426,
final_local_scores=0.78505594, pem=0.0, similarity=27.517347]
34: 31737 Supreme_Court_of_the_United_States
[pem_pos=0, final_scores=0.3256827,
final_local_scores=0.3457474, pem=-0.9940822,
similarity=19.813478]
35: 31737 Supreme_Court_of_the_United_States
[pem_pos=0, final_scores=0.39531678,
final_local_scores=0.42181212, pem=-0.9940822,
similarity=22.083403]
36: 32070 Republican_Party_(United_States)
[pem_pos=0, final_scores=0.437208,
final_local_scores=0.40849715, pem=-0.11427376,
similarity=16.901522]
37: 33057 White_House
[pem_pos=0, final_scores=0.48926812,
final_local_scores=0.478611, pem=-0.0056634876,
similarity=18.403217]
38: 31756 United_States_Congress
[pem_pos=0, final_scores=0.4122308,
final_local_scores=0.40105504, pem=-0.3349529,
similarity=17.879522]
39: 58353 Medicare_(United_States)
[pem_pos=0, final_scores=0.4265775,
final_local_scores=0.48308128, pem=0.0, similarity=18.505821]
40: 32070 Republican_Party_(United_States)
[pem_pos=0, final_scores=0.44955534,
final_local_scores=0.42198467, pem=-0.11427376,
similarity=17.304016]
41: 33127 Wisconsin
[pem_pos=0, final_scores=0.56836987,
final_local_scores=0.6169415, pem=-0.21268336,
similarity=23.657084]
42: 3434750 United_States
[pem_pos=0, final_scores=0.588787,
final_local_scores=0.6370457, pem=-0.78483206,
similarity=27.368465]
43: 3356 Bill_Clinton
[pem_pos=0, final_scores=0.60891855,
final_local_scores=0.5681012, pem=0.0, similarity=21.042986]
44: 55846 Bob_Dole
[pem_pos=0, final_scores=0.9397678,
final_local_scores=0.9413803, pem=0.0, similarity=32.18238]
45: 5042706 Al_Gore
[pem_pos=0, final_scores=0.7858897,
final_local_scores=0.7546308, pem=0.0, similarity=26.609402]
46: 55846 Bob_Dole
[pem_pos=0, final_scores=0.9322317,
final_local_scores=0.9331483, pem=0.0, similarity=31.936722]
47: 3356 Bill_Clinton
[pem_pos=0, final_scores=0.8537852,
final_local_scores=0.8364369, pem=0.0, similarity=29.05066]

```
48: 33127 Wisconsin
    [pem_pos=0, final_scores=0.57099307,
final_local_scores=0.61980695, pem=-0.21268336,
similarity=23.742594]
49: 5042706 Al_Gore
    [pem_pos=0, final_scores=0.9251266,
final_local_scores=0.90672606, pem=0.0, similarity=31.148226]
50: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.9570841,
final_local_scores=0.96029574, pem=0.0, similarity=32.746853]
51: 5042706 Al_Gore
    [pem_pos=0, final_scores=0.76184124,
final_local_scores=0.7283616, pem=0.0, similarity=25.825476]
52: 33127 Wisconsin
    [pem_pos=0, final_scores=0.48582083,
final_local_scores=0.5267692, pem=-0.21268336,
similarity=20.966162]
53: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.85916597,
final_local_scores=0.8423146, pem=0.0, similarity=29.22606]
54: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.49054152,
final_local_scores=0.46675587, pem=-0.1310191,
similarity=18.731142]
55: 31756 United_States_Congress
    [pem_pos=0, final_scores=0.45394808,
final_local_scores=0.44662488, pem=-0.3349529,
similarity=19.239416]
56: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.44112748,
final_local_scores=0.41277856, pem=-0.11427376,
similarity=17.029287]
57: 5042706 Al_Gore
    [pem_pos=0, final_scores=0.85042435,
final_local_scores=0.8251253, pem=0.0, similarity=28.713097]
```

Sample: b'485_Actor'

Actor [1 1 Reeve 1] highlights [2 2 Democrats 2] ' first night . [3 3 CHICAGO 3] 1996-08-26 The highlights of the first day of the [4 4 Democratic 4] convention on Monday were expected to feature a mix of party leaders and people who have overcome adversity . Among the latter were gun control advocate [5 5 Sarah Brady 5] and actor [6 6 Christopher Reeve 6] and the politicians included 1 Rep . [7 7 Richard Gephardt 7] and Sen . [8 8 Thomas Daschle 8] . Here are thumbnail profiles of the convention 's key Monday speakers . [9 9 Sarah Brady 9] -- The nation 's toughest gun control law is named after [10 10 Ronald Reagan 10] 's press secretary [11 11 James Brady 11] but it was his wife who was the major force behind its passage . As head of Handgun Control Inc. , [12 12 Sarah Brady 12] , 54 , campaigned nonstop for tough gun control in the years following the shooting of her husband and then

President [13 13 Reagan 13] in 1981 . Her reward was the passage in 1993 of the " Brady Bill " which requires a mandatory five-day waiting period for purchase of handguns and also mandates background checks for would-be gun purchasers . [14 14 Reagan 14] recovered fully from his wounds but [15 15 Brady 15] , who was close to death after being shot by [16 16 John Hinckley Jr 16] . , suffered serious brain damage . [17 17 Sarah 17] was [18 18 Brady 18] 's second wife and they have a son , James Scott Brady Jr . Before the assassination attempt , she had worked for [19 19 Republican 19] congressmen and for the [20 20 Republican Party 20] . [21 21 Christopher Reeve 21] -- [22 22 Reeve 22] was best known for playing the comic book hero 2 Superman in four movies but his greatest heroics came in real life . [23 23 Reeve 23] , an accomplished rider who owned several horses , suffered multiple injuries including two shattered neck vertebrae when he was thrown from his horse at an equestrian event in [24 24 Culpepper 24] , [25 25 Virginia 25] , on May 27 , 1995 . Almost entirely paralyzed , [26 26 Reeve 26] underwent extensive surgery to fuse the vertebrae to the base of the skull and prevent any further damage to his spine . That allowed him to be moved to a semi-upright position . Over time he regained the power of speech , so much so that he was asked to address the opening night of the [27 27 Democratic National Convention 27] . [28 28 Reeve 28] , 43 , was classically trained as an actor but became the prototypical handsome leading man . He performed in summer stock and soap operas before being plucked as an almost unknown to play the lead in " 3 Superman " and three sequels . [29 29 Richard Gephardt 29] -- [30 30 Gephardt 30] , House [31 31 Democratic 31] leader , is a politician with a " [32 32 Mr Clean 32] " reputation who sought the presidency eight years ago and is widely believed to still have ambitions for the job . [33 33 Gephardt 33] , 55 , the son of a milkman from a working class district of [34 34 St. Louis 34] , is a consummate congressional insider , sufficiently skilled in compromise and the ways of the legislature to manage the often-unruly House [35 35 Democrats 35] . A former lawyer , he was in the front line of President [36 36 Bill Clinton 36] 's battle with the [37 37 Republican-led 37] [38 38 Congress 38] over the budget but has opposed the president 's decision to sign the Republican-written welfare reform bill . He advocated tough action against foreign countries to cut [39 39 U.S. 39] trade deficits but sometimes been out of step with the party 's liberal wing . He was an opponent of abortion until 1986 and voted for President [40 40 Ronald Reagan 40] 's big tax cut bill . [41 41 Gephardt 41] , a red-haired square-jawed man , is a less than fiery orator . His 1988 bid for the [42 42 Democratic 42] nomination only took off after he recreated himself as a firebreathing reformer of the establishment , standing up for blue collar workers and farmers . [43 43 Tom Daschle 43] -- [44 44 Daschle 44] , 48 , was largely unknown

outside [45 45 Washington 45] and his state of [46 46 South Dakota 46] when he surprisingly beat more prominent rivals to become Senate [47 47 Democratic 47] leader after the party lost its majority to the [48 48 Republicans 48] in 1994 . A mild-mannered and youthful man who rose rapidly after entering the Senate in 1987 , he presented himself as a [49 49 Midwest 49] moderate , as a [50 50 Democratic 50] winner in a [51 51 Republican 51] state able to unite Senate factions . Concern that he might be steamrollered by the vastly more experienced [52 52 Republican 52] leader [53 53 Bob Dole 53] was dispelled when he showed a tough edge , outmaneuvering [54 54 Dole 54] in a wrangle over scrapping a gas tax and raising the federal minimum wage . In his early Senate years he was seen as a " prairie populist " , working on legislation protecting farmers ' prices and also on compensating veterans sickened by [55 55 Agent Orange 55] defoliant spraying in the [56 56 Vietnam War 56] . [57 57 Daschle 57] was a key player in President [58 58 Bill Clinton 58] 's failed attempt at sweeping healthcare changes but when he became minority leader he declared he would be no water-carrier for the [59 59 White House 59] . He has made clear he opposed [60 60 Clinton 60] 's signing of the Republican-initiated welfare reform bill . He has spent his adult life in politics , coming to [61 61 Congress 61] as an aide in 1972 after three years in the [62 62 Air Force 62] and then being elected to the House of Representatives himself in 1978 .

1: 73626 Christopher_Reeve
2: 5043544 Democratic_Party_(United_States)
3: 6886 Chicago
4: 5043544 Democratic_Party_(United_States)
5: 1850769 Sarah_Brady
6: 73626 Christopher_Reeve
7: 165912 Dick_Gephardt
8: 30930 Tom_Daschle
9: 1850769 Sarah_Brady
10: 25433 Ronald_Reagan
11: 214323 James_Brady
12: 1850769 Sarah_Brady
13: 25433 Ronald_Reagan
14: 25433 Ronald_Reagan
15: 214323 James_Brady
16: 222294 John_Hinckley,_Jr.
17: 1850769 Sarah_Brady
18: 214323 James_Brady
19: 32070 Republican_Party_(United_States)
20: 32070 Republican_Party_(United_States)
21: 73626 Christopher_Reeve
22: 73626 Christopher_Reeve
23: 73626 Christopher_Reeve
24: 137552 Culpeper,_Virginia
25: 32432 Virginia
26: 73626 Christopher_Reeve

```
27: 468102 Democratic_National_Convention
28: 73626 Christopher_Reeve
29: 165912 Dick_Gephardt
30: 165912 Dick_Gephardt
31: 5043544 Democratic_Party_(United_States)
32: 736403 Mr._Clean
33: 165912 Dick_Gephardt
34: 27687 St._Louis
35: 5043544 Democratic_Party_(United_States)
36: 3356 Bill_Clinton
37: 32070 Republican_Party_(United_States)
38: 31756 United_States_Congress
39: 3434750 United_States
40: 25433 Ronald_Reagan
41: 165912 Dick_Gephardt
42: 5043544 Democratic_Party_(United_States)
43: 30930 Tom_Daschle
44: 30930 Tom_Daschle
45: 108956 Washington,_D.C.
46: 26746 South_Dakota
47: 5043544 Democratic_Party_(United_States)
48: 32070 Republican_Party_(United_States)
49: 104697 Midwestern_United_States
50: 5043544 Democratic_Party_(United_States)
51: 32070 Republican_Party_(United_States)
52: 32070 Republican_Party_(United_States)
53: 55846 Bob_Dole
54: 55846 Bob_Dole
55: 2547 Agent_Orange
56: 32611 Vietnam_War
57: 30930 Tom_Daschle
58: 3356 Bill_Clinton
59: 33057 White_House
60: 3356 Bill_Clinton
61: 31756 United_States_Congress
62: 32090 United_States_Air_Force
1: 32070 Republican_Party_(United_States)
    [winner: 32070 Republican_Party_(United_States),
pem_pos=0, final_scores=-0.08910377,
final_local_scores=-0.17393449, pem=-0.9360621,
similarity=3.989612|
     gt: 32070 Republican_Party_(United_States), pem_pos=0,
final_scores=-0.08910377, final_local_scores=-0.17393449,
pem=-0.9360621, similarity=3.989612 ]
    [winner: 4721653 Zyphe_language, pem_pos=12,
final_scores=-1.4818922, final_local_scores=-1.449374,
pem=-4.4188404, similarity=-15.132155|
     RECALL MISS: 32070 Republican_Party_(United_States)]
2: 362719 Superman_(1978_film)
    [winner: 28381 Superman, pem_pos=0,
final_scores=-0.07660152, final_local_scores=-0.06906763,
pem=-0.21335615, similarity=3.1888611|
```

RECALL MISS: 362719 Superman_(1978_film)]
3: 362719 Superman_(1978_film)
[winner: 28381 Superman, pem_pos=0,
final_scores=-0.18142775, final_local_scores=-0.18357435,
pem=-0.21335615, similarity=-0.22824809]
RECALL MISS: 362719 Superman_(1978_film)]
1: 73626 Christopher_Reeve
[pem_pos=0, final_scores=1.1335368,
final_local_scores=1.2150662, pem=0.0, similarity=40.34971]
2: 5043544 Democratic_Party_(United_States)
[pem_pos=0, final_scores=0.48699838,
final_local_scores=0.45789903, pem=-0.18007085,
similarity=18.733587]
3: 6886 Chicago
[pem_pos=0, final_scores=0.6118236,
final_local_scores=0.63437444, pem=-0.20658566,
similarity=24.144156]
4: 5043544 Democratic_Party_(United_States)
[pem_pos=1, final_scores=0.4467405,
final_local_scores=0.41392344, pem=-0.076093495,
similarity=16.855824]
5: 1850769 Sarah_Brady
[pem_pos=0, final_scores=0.4041654,
final_local_scores=0.46469557, pem=0.0, similarity=17.957155]
6: 73626 Christopher_Reeve
[pem_pos=0, final_scores=0.53029996,
final_local_scores=0.5528788, pem=0.0, similarity=20.588718]
7: 165912 Dick_Gephardt
[pem_pos=0, final_scores=0.60946316,
final_local_scores=0.60684884, pem=0.0, similarity=22.199293]
8: 30930 Tom_Daschle
[pem_pos=0, final_scores=0.5285441,
final_local_scores=0.5419063, pem=0.0, similarity=20.261278]
9: 1850769 Sarah_Brady
[pem_pos=0, final_scores=0.46403044,
final_local_scores=0.5300891, pem=0.0, similarity=19.908628]
10: 25433 Ronald_Reagan
[pem_pos=0, final_scores=0.39307088,
final_local_scores=0.34124213, pem=0.0, similarity=14.273059]
11: 214323 James_Brady
[pem_pos=0, final_scores=0.48165834,
final_local_scores=0.53426564, pem=0.0, similarity=20.033266]
12: 1850769 Sarah_Brady
[pem_pos=0, final_scores=0.6019925,
final_local_scores=0.6807916, pem=0.0, similarity=24.405893]
13: 25433 Ronald_Reagan
[pem_pos=0, final_scores=0.42037815,
final_local_scores=0.37246776, pem=0.0, similarity=15.204893]
14: 25433 Ronald_Reagan
[pem_pos=0, final_scores=0.43545043,
final_local_scores=0.38893193, pem=0.0, similarity=15.696218]
15: 214323 James_Brady

```
[pem_pos=1, final_scores=0.38411522,
final_local_scores=0.4277146, pem=0.0, similarity=16.853569]
16: 222294 John_Hinckley,_Jr.
    [pem_pos=0, final_scores=0.49120468,
final_local_scores=0.5437491, pem=0.0, similarity=20.31627]
17: 1850769 Sarah_Brady
    [pem_pos=0, final_scores=0.5488977,
final_local_scores=0.6227936, pem=0.0, similarity=22.675117]
18: 214323 James_Brady
    [pem_pos=1, final_scores=0.40198845,
final_local_scores=0.4472384, pem=0.0, similarity=17.4362]
19: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.46442467,
final_local_scores=0.43245298, pem=-0.1310191,
similarity=17.707476]
20: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.49379587,
final_local_scores=0.46453655, pem=-0.16706309,
similarity=18.860926]
21: 73626 Christopher_Reeve
    [pem_pos=0, final_scores=0.5763057,
final_local_scores=0.603133, pem=0.0, similarity=22.088406]
22: 73626 Christopher_Reeve
    [pem_pos=0, final_scores=1.1261513,
final_local_scores=1.2069988, pem=0.0, similarity=40.108963]
23: 73626 Christopher_Reeve
    [pem_pos=0, final_scores=1.3297298,
final_local_scores=1.4293773, pem=0.0, similarity=46.745182]
24: 137552 Culpeper,_Virginia
    [pem_pos=1, final_scores=0.38878936,
final_local_scores=0.44847554, pem=-1.2981198,
similarity=24.532495]
25: 32432 Virginia
    [pem_pos=0, final_scores=0.63708764,
final_local_scores=0.6712203, pem=-0.2023721,
similarity=25.220797]
26: 73626 Christopher_Reeve
    [pem_pos=0, final_scores=1.2794375,
final_local_scores=1.3744407, pem=0.0, similarity=45.105762]
27: 468102 Democratic_National_Convention
    [pem_pos=0, final_scores=0.36053467,
final_local_scores=0.33673126, pem=-0.5317093,
similarity=17.029963]
28: 73626 Christopher_Reeve
    [pem_pos=0, final_scores=1.2400467,
final_local_scores=1.3314121, pem=0.0, similarity=43.821705]
29: 165912 Dick_Gephardt
    [pem_pos=0, final_scores=0.72310376,
final_local_scores=0.7309839, pem=0.0, similarity=25.90373]
30: 165912 Dick_Gephardt
    [pem_pos=0, final_scores=0.5417407,
final_local_scores=0.5328724, pem=0.0, similarity=19.991688]
```

```
31: 5043544 Democratic_Party_(United_States)
    [pem_pos=1, final_scores=0.44678146,
final_local_scores=0.4139682, pem=-0.076093495,
similarity=16.857159]
32: 736403 Mr._Clean
    [pem_pos=0, final_scores=0.42457718,
final_local_scores=0.5475121, pem=0.0, similarity=20.428566]
33: 165912 Dick_Gephardt
    [pem_pos=0, final_scores=0.4761706,
final_local_scores=0.46124697, pem=0.0, similarity=17.854242]
34: 27687 St._Louis
    [pem_pos=0, final_scores=0.36914086,
final_local_scores=0.38215792, pem=-0.28531665,
similarity=17.045664]
35: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.4663872,
final_local_scores=0.4353845, pem=-0.18007085,
similarity=18.061708]
36: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.5524116,
final_local_scores=0.52352834, pem=0.0, similarity=19.712843]
37: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.4507926,
final_local_scores=0.41756207, pem=0.0, similarity=16.550598]
38: 31756 United_States_Congress
    [pem_pos=0, final_scores=0.46133906,
final_local_scores=0.45115048, pem=-0.3349529,
similarity=19.37447]
39: 3434750 United_States
    [pem_pos=0, final_scores=0.49766797,
final_local_scores=0.5204287, pem=-0.14075686,
similarity=20.385801]
40: 25433 Ronald_Reagan
    [pem_pos=0, final_scores=0.3621983,
final_local_scores=0.30751848, pem=0.0, similarity=13.266678]
41: 165912 Dick_Gephardt
    [pem_pos=0, final_scores=0.38401926,
final_local_scores=0.3605856, pem=0.0, similarity=14.8503065]
42: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.46990913,
final_local_scores=0.43923163, pem=-0.076093495,
similarity=17.61107]
43: 30930 Tom_Daschle
    [pem_pos=0, final_scores=0.7417421,
final_local_scores=0.7747927, pem=0.0, similarity=27.211071]
44: 30930 Tom_Daschle
    [pem_pos=0, final_scores=0.5977823,
final_local_scores=0.6175385, pem=0.0, similarity=22.518293]
45: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.534137,
final_local_scores=0.52748334, pem=-1.7496481,
similarity=29.345726]
```

```
46: 26746 South_Dakota
    [pem_pos=0, final_scores=0.35045475,
final_local_scores=0.3676831, pem=-0.23758878,
similarity=16.354156]
47: 5043544 Democratic_Party_(United_States)
    [pem_pos=1, final_scores=0.46727604,
final_local_scores=0.43635535, pem=-0.076093495,
similarity=17.525236]
48: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.4200012,
final_local_scores=0.3839271, pem=-0.11427376,
similarity=16.168303]
49: 104697 Midwestern_United_States
    [pem_pos=0, final_scores=0.37112892,
final_local_scores=0.41142648, pem=-0.031208755,
similarity=16.537218]
50: 5043544 Democratic_Party_(United_States)
    [pem_pos=1, final_scores=0.45291483,
final_local_scores=0.42066795, pem=-0.076093495,
similarity=17.057093]
51: 32070 Republican_Party_(United_States)
    [pem_pos=1, final_scores=0.4256444,
final_local_scores=0.39009142, pem=-0.1310191,
similarity=16.443321]
52: 32070 Republican_Party_(United_States)
    [pem_pos=1, final_scores=0.4614529,
final_local_scores=0.4292068, pem=-0.1310191,
similarity=17.610601]
53: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.771663,
final_local_scores=0.77516395, pem=0.0, similarity=27.22215]
54: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.730179,
final_local_scores=0.7318961, pem=0.0, similarity=25.930952]
55: 2547 Agent_Orange
    [pem_pos=0, final_scores=0.37401545,
final_local_scores=0.4627542, pem=-0.029288419,
similarity=18.058496]
56: 32611 Vietnam_War
    [pem_pos=0, final_scores=0.40259522,
final_local_scores=0.4168408, pem=0.0, similarity=16.529074]
57: 30930 Tom_Daschle
    [pem_pos=0, final_scores=0.5958453,
final_local_scores=0.61542267, pem=0.0, similarity=22.455153]
58: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.5455614,
final_local_scores=0.51604545, pem=0.0, similarity=19.489538]
59: 33057 White_House
    [pem_pos=0, final_scores=0.46425885,
final_local_scores=0.4465853, pem=-0.0056634876,
similarity=17.447508]
60: 3356 Bill_Clinton
```

```
[pem_pos=0, final_scores=0.8395272,
final_local_scores=0.83882296, pem=0.0, similarity=29.121862]
61: 31756 United_States_Congress
    [pem_pos=0, final_scores=0.3724882,
final_local_scores=0.35409442, pem=-0.3349529,
similarity=16.478123]
62: 32090 United_States_Air_Force
    [pem_pos=0, final_scores=0.3528723,
final_local_scores=0.3867464, pem=-0.8379956,
similarity=20.188147]
```

Sample: b'486_U.S.'

```
[1 1 U.S. 1] boy , 13 , accused of murdering adoptive
mother . [2 2 DALLAS 2] 1996-08-26 A 13-year-old [3 3 Dallas 3
] boy has been charged with murdering his adoptive mother
over the weekend , police said on Monday . Margaret
McCullough , 55 , was found dead in her home on Saturday with
gunshot wounds to the head . Police , who at first thought
her son had been kidnapped , found him on Sunday with a
friend in his mother 's car in [4 4 Oklahoma 4] and arrested
him . A shotgun and a .357 handgun were found in the car . A
police spokesman said the boy was being questioned . " They
are talking to him now about the motive and everything else .
" McCullough had adopted the boy , who was the grandson of
her late husband , shortly after his birth but neighbours
said they often had loud arguments .
```

```
1: 3434750 United_States
```

```
2: 53838 Dallas
```

```
3: 53838 Dallas
```

```
4: 22489 Oklahoma
```

```
1: 3434750 United_States
```

```
    [pem_pos=0, final_scores=0.46764404,
final_local_scores=0.45434076, pem=-0.14075686,
similarity=18.413605]
```

```
2: 53838 Dallas
```

```
    [pem_pos=0, final_scores=0.49101013,
final_local_scores=0.43824828, pem=-0.4621424,
similarity=19.681118]
```

```
3: 53838 Dallas
```

```
    [pem_pos=0, final_scores=0.58328944,
final_local_scores=0.5390494, pem=-0.4621424,
similarity=22.689224]
```

```
4: 22489 Oklahoma
```

```
    [pem_pos=0, final_scores=0.6582082,
final_local_scores=0.65803885, pem=-0.35266355,
similarity=25.644743]
```

Sample: b'487_Pakistani'

```
[1 1 Pakistani 1] bourse to use new recomposed index . [2 2 K
ARACHI 2] , [3 3 Pakistan 3] 1996-08-26 The [4 4 Karachi Stock
Exchange 4] ( [5 5 KSE 5] ) said on Monday it would introduce
a new recomposed [6 6 KSE-100 6] index on September 10 . A [7
```

7 KSE 7] statement said the recomposed index would be more representative , capturing market capitalisation of 82.3 percent , up from 79.9 percent previously . -- [8 8 Karachi 8] newsroom 9221-5685192

1: 23235 Pakistan
2: 17123 Karachi
3: 23235 Pakistan
4: 1466645 Karachi_Stock_Exchange
5: 1466645 Karachi_Stock_Exchange
6: 1965945 KSE_100_Index
7: 1466645 Karachi_Stock_Exchange
8: 17123 Karachi
1: 23235 Pakistan
[pem_pos=0, final_scores=0.48156828, final_local_scores=0.4786561, pem=-0.49826378, similarity=21.0834]
2: 17123 Karachi
[pem_pos=0, final_scores=0.4233957, final_local_scores=0.411012, pem=0.0, similarity=16.355131]
3: 23235 Pakistan
[pem_pos=0, final_scores=0.627729, final_local_scores=0.63831455, pem=-0.21944217, similarity=24.331652]
4: 1466645 Karachi_Stock_Exchange
[pem_pos=0, final_scores=0.6759779, final_local_scores=0.6896008, pem=0.0, similarity=24.668777]
5: 1466645 Karachi_Stock_Exchange
[pem_pos=1, final_scores=0.35615605, final_local_scores=0.340244, pem=-0.8485613, similarity=18.857882]
6: 1965945 KSE_100_Index
[pem_pos=0, final_scores=0.37584633, final_local_scores=0.44516933, pem=0.0, similarity=17.374454]
7: 1466645 Karachi_Stock_Exchange
[pem_pos=1, final_scores=0.35354143, final_local_scores=0.33738792, pem=-0.8485613, similarity=18.77265]
8: 17123 Karachi
[pem_pos=0, final_scores=0.41289538, final_local_scores=0.39954203, pem=0.0, similarity=16.012844]

Sample: b'488_NWE'

NWE oil products mixed , holiday dulls trade . [1 1 LONDON 1] 1996-08-26 NWE oil products were mixed on Monday but markets were becalmed because of a public holiday in the [2 2 United Kingdom 2] , traders said . An explosion at [3 3 Repsol 3] 's [4 4 Puertollano 4] refinery , which killed four workers , had not affected output of oil products , an official said . " The plant is functioning as usual , " Jose Manuel Prieto , director of personnel , told [5 5 Spanish 5] state television . Gasoline prices were notionally unchanged from Friday despite sagging NYMEX numbers , with the arbitrage

window to the [6 6 U.S. 6] considered closed for the moment . Eurograde barges were offered at \$ 207 fob ARA for Amsterdam-Rotterdam barrels , and at \$ 206 for full ARA material . " There is no market at the moment , " one [7 7 Rotterdam 7] trader said . " Maybe sentiment is a little bit weaker but prices have not changed . " Outright gas oil prices were notionally softer as the NYMEX heating oil contract headed lower , and following news that the Indian Oil Corp ([8 8 IOC 8]) had issued a tender to buy only 120,000 tonnes of high speed diesel for October . [9 9 Asian 9] traders had earlier expected an [10 10 IOC 10] tender for around 400,000 tonnes . ARA gas oil barges were quiet although one trader said he had seen offers at \$ 1 a tonne over September IPE for prompt barrels , while [11 11 Antwerp 11] material was available for 0-50 cents over September . The [12 12 IOC 12] tender had " a bearish impact , but not a great impact , " one [13 13 German 13] player said . Fuel oil markets were also listless . Offers were around a dollar a tonne higher at \$ 102 fob ARA but bids were scarce . -- Nicholas Shaxson , [14 14 London 14] newsroom +44 171 542 8167

1: 17867 London
2: 31717 United_Kingdom
3: 1076420 Repsol_S.A.
4: 3704829 Puertollano
5: 26667 Spain
6: 3434750 United_States
7: 26049 Rotterdam
8: 7971634 Indian_Oil_Corporation
9: 689 Asia
10: 7971634 Indian_Oil_Corporation
11: 32149462 Antwerp
12: 7971634 Indian_Oil_Corporation
13: 11867 Germany
14: 17867 London
1: 17867 London
[pem_pos=0, final_scores=0.6907343,
final_local_scores=0.69407684, pem=-0.09086737,
similarity=25.296501]
2: 31717 United_Kingdom
[pem_pos=0, final_scores=0.46387362,
final_local_scores=0.44463515, pem=-0.09933331,
similarity=17.898703]
3: 1076420 Repsol_S.A.
[pem_pos=0, final_scores=0.5699497,
final_local_scores=0.6534316, pem=0.0, similarity=23.589415]
4: 3704829 Puertollano
[pem_pos=0, final_scores=0.23098631,
final_local_scores=0.333123, pem=0.0, similarity=14.030767]
5: 26667 Spain
[pem_pos=1, final_scores=0.60308576,
final_local_scores=0.62507945, pem=-1.2425895,
similarity=29.500725]

```
6: 3434750 United_States
    [pem_pos=0, final_scores=0.47652513,
final_local_scores=0.46630388, pem=-0.14075686,
similarity=18.770607]
7: 26049 Rotterdam
    [pem_pos=0, final_scores=0.41897196,
final_local_scores=0.42371184, pem=-0.145035,
similarity=17.522842]
8: 7971634 Indian_Oil_Corporation
    [pem_pos=2, final_scores=0.70334417,
final_local_scores=0.79367036, pem=-1.36746,
similarity=35.21088]
9: 689 Asia
    [pem_pos=1, final_scores=0.38522136,
final_local_scores=0.4080031, pem=-1.9623682,
similarity=26.937002]
10: 7971634 Indian_Oil_Corporation
    [pem_pos=2, final_scores=0.69527334,
final_local_scores=0.78485423, pem=-1.36746,
similarity=34.94779]
11: 32149462 Antwerp
    [pem_pos=0, final_scores=0.40693182,
final_local_scores=0.41870052, pem=-0.11904648,
similarity=17.231966]
12: 7971634 Indian_Oil_Corporation
    [pem_pos=2, final_scores=0.61503124,
final_local_scores=0.6972019, pem=-1.36746,
similarity=32.33207]
13: 11867 Germany
    [pem_pos=1, final_scores=0.632201,
final_local_scores=0.63930064, pem=-1.0465823,
similarity=28.859198]
14: 17867 London
    [pem_pos=0, final_scores=0.51776785,
final_local_scores=0.5051372, pem=-0.09086737,
similarity=19.658165]
```

Sample: b'489_Prairies'

[1 1 Prairies 1] saw no frost Monday , none rest of week . [2
2 WINNIPEG 2] 1996-08-26 [3 3 Canada 3] 's [4 4 Prairies 4]
saw no frost on Monday morning and none was expected anywhere
on the grainbelt until late in the [5 5 Labour Day 5] long
weekend , [6 6 Environment Canada 6] said . " Apparently , we
're home free for the rest of the week . We 're not calling
for any frost until after the weekend when it starts to cool
off in northwestern [7 7 Alberta 7] after the weekend
probably Monday or Tuesday , " meteorologist Gerald Machnee
told [8 8 Reuters 8] . [9 9 Sprague 9] , Manitoba , on the [10
10 Minnesota 10] border was the cold spot of the [11 11 Prair
ies 11] Monday morning at 4.0 [12 12 Celsius 12] (39.2 F) .
Temperatures at ground level can be 2.0 to 5.0 [13 13 Celsius
13] lower than at chest level depending on windspeed , sky

conditions and ground surface moisture . Freezing occurs at 0 [14 14 Celsius 14] (32.0 F) . [15 15 North Battleford 15] , Sask . , reported a low of 5.0 [16 16 Celsius 16] (41.0 F) and [17 17 Grande Prairie 17] , Alta . , in the [18 18 Peace River Valley 18] reported 7.0 [19 19 Celsius 19] (44.6 F) . Machnee dismissed talk of frost Wednesday by proponents of the " full moon , frost soon " school of thought . Lows on August 28 across the [20 20 Prairies 20] should range from 8.0 to 12.0 [21 21 Celsius 21] with highs around 30.0 [22 22 Celsius 22] . -- Gilbert Le Gras 204 947 3548

1: 437251 Canadian_Prairies
2: 100730 Winnipeg
3: 5042916 Canada
4: 437251 Canadian_Prairies
5: 18051 Labour_Day
6: 272329 Environment_Canada
7: 717 Alberta
8: 18998750 Reuters
9: 16593248 Sprague,_Manitoba
10: 19590 Minnesota
11: 437251 Canadian_Prairies
12: 19593040 Celsius
13: 19593040 Celsius
14: 19593040 Celsius
15: 331545 North_Battleford
16: 19593040 Celsius
17: 329344 Grande_Prairie
18: 749287 Peace_River
19: 19593040 Celsius
20: 437251 Canadian_Prairies
21: 19593040 Celsius
22: 19593040 Celsius
1: 437251 Canadian_Prairies
 [pem_pos=0, final_scores=0.45794994,
final_local_scores=0.4395427, pem=0.0, similarity=17.206545]
2: 100730 Winnipeg
 [pem_pos=0, final_scores=0.3963638,
final_local_scores=0.4103139, pem=-0.123312846,
similarity=17.004892]
3: 5042916 Canada
 [pem_pos=0, final_scores=0.6507753,
final_local_scores=0.6664153, pem=-0.11662793,
similarity=24.611115]
4: 437251 Canadian_Prairies
 [pem_pos=0, final_scores=0.41546685,
final_local_scores=0.39313632, pem=0.0, similarity=15.821686]
5: 18051 Labour_Day
 [pem_pos=0, final_scores=0.3068264,
final_local_scores=0.38942087, pem=0.0, similarity=15.710808]
6: 272329 Environment_Canada
 [pem_pos=0, final_scores=0.47201556,
final_local_scores=0.51271063, pem=0.0, similarity=19.39002]

7: 717 Alberta
[pem_pos=0, final_scores=0.5905785,
final_local_scores=0.59423995, pem=-0.06637208,
similarity=22.18396]
8: 18998750 Reuters
[pem_pos=0, final_scores=0.33506656,
final_local_scores=0.4371689, pem=0.0, similarity=17.135704]
9: 16593248 Sprague, Manitoba
[pem_pos=9, final_scores=0.3533554,
final_local_scores=0.5034877, pem=-3.2257934,
similarity=36.657154]
10: 19590 Minnesota
[pem_pos=0, final_scores=0.4038537,
final_local_scores=0.42821413, pem=-0.21388707,
similarity=18.031628]
11: 437251 Canadian_Prairies
[pem_pos=0, final_scores=0.47808444,
final_local_scores=0.4615366, pem=0.0, similarity=17.862885]
12: 19593040 Celsius
[pem_pos=0, final_scores=0.7198381,
final_local_scores=0.761612, pem=0.0, similarity=26.817734]
13: 19593040 Celsius
[pem_pos=0, final_scores=0.6790535,
final_local_scores=0.7170609, pem=0.0, similarity=25.488241]
14: 19593040 Celsius
[pem_pos=0, final_scores=0.66938657,
final_local_scores=0.7065013, pem=0.0, similarity=25.17312]
15: 331545 North_Battleford
[pem_pos=0, final_scores=0.49375266,
final_local_scores=0.54823494, pem=0.0, similarity=20.450136]
16: 19593040 Celsius
[pem_pos=0, final_scores=0.7181451,
final_local_scores=0.7597626, pem=0.0, similarity=26.762545]
17: 329344 Grande_Prairie
[pem_pos=0, final_scores=0.3530661,
final_local_scores=0.3833602, pem=0.0, similarity=15.529945]
18: 749287 Peace_River
[pem_pos=0, final_scores=0.42195237,
final_local_scores=0.47121632, pem=0.0, similarity=18.151749]
19: 19593040 Celsius
[pem_pos=0, final_scores=0.7186263,
final_local_scores=0.7602883, pem=0.0, similarity=26.778233]
20: 437251 Canadian_Prairies
[pem_pos=0, final_scores=0.3742872,
final_local_scores=0.34815386, pem=0.0, similarity=14.479319]
21: 19593040 Celsius
[pem_pos=0, final_scores=0.67383605,
final_local_scores=0.7113617, pem=0.0, similarity=25.318165]
22: 19593040 Celsius
[pem_pos=0, final_scores=0.64691573,
final_local_scores=0.6819554, pem=0.0, similarity=24.440622]

Sample: b'490_Brush'

[1 1 Brush Wellman 1] comments on beryllium lawsuits . [2 2 C LEVELAND 2] 1996-08-26 Brush Wellman Inc said Monday that 10 of 24 lawsuits involving chronic beryllium disease have been dismissed since July 1 . The leading [3 3 U.S. 3] beryllium producer said in a conference call it has traditionally been pro-active regarding the workplace disease , a lung ailment which can affects a small percent of people whose immune systems are susceptible . Of the 14 remaining suits , 10 were filed by employees of industrial [4 4 Brush Wellman 4] customers and [5 5 Brush Wellman 5] liability in such suits is typically covered by insurance , Timothy Reid , vice president of corporate communications , said on the call . The company was responding to an article in Sunday 's [6 6 New York Times 6] . He said the article largely reiterated information about the suits and the disease which had previously been made public via Securities and Exchange Commission filings and annual reports . " [7 7 Brush Wellman 7] has been a leader in dealing with health and safety issues (related to chronic Beryllium disease) for nearly 50 years , " he said . " We have a record of going beyond regulatory requirements ... and we consistently share the most current information available ... with customers and employees . " The customer employee suits were filed in 1990-95 , he said . A class action filed by a former [8 8 Brush Wellman 8] employee in April 1996 was dismissed in July , he said . The company is " vigorously defending " the remaining four suits filed by former and current [9 9 Brush Wellman 9] employees , he said . After a delayed opening , the stock was off 1-1/2 to 18-7/8 . -- Cleveland Newsdesk 216-579-0077

1: 8372532 Brush_Engineered_Materials
2: 5951 Cleveland
3: 3434750 United_States
4: 8372532 Brush_Engineered_Materials
5: 8372532 Brush_Engineered_Materials
6: 30680 The_New_York_Times
7: 8372532 Brush_Engineered_Materials
8: 8372532 Brush_Engineered_Materials
9: 8372532 Brush_Engineered_Materials
1: 8372532 Brush_Engineered_Materials
 [pem_pos=0, final_scores=0.46690434,
final_local_scores=0.41648406, pem=0.0, similarity=16.518429]
2: 5951 Cleveland
 [pem_pos=0, final_scores=0.3637027,
final_local_scores=0.48723304, pem=-0.57246673,
similarity=21.742882]
3: 3434750 United_States
 [pem_pos=0, final_scores=0.38073498,
final_local_scores=0.51815015, pem=-0.14075686,
similarity=20.317804]
4: 8372532 Brush_Engineered_Materials

```
[pem_pos=0, final_scores=0.49696624,
final_local_scores=0.44932216, pem=0.0, similarity=17.498383]
5: 8372532 Brush_Engineered_Materials
[pem_pos=0, final_scores=0.51738995,
final_local_scores=0.47163194, pem=0.0, similarity=18.16415]
6: 30680 The_New_York_Times
[pem_pos=0, final_scores=0.2691731,
final_local_scores=0.4260648, pem=0.0, similarity=16.804337]
7: 8372532 Brush_Engineered_Materials
[pem_pos=0, final_scores=0.5938545,
final_local_scores=0.55515784, pem=0.0, similarity=20.65673]
8: 8372532 Brush_Engineered_Materials
[pem_pos=0, final_scores=0.57133526,
final_local_scores=0.530559, pem=0.0, similarity=19.922651]
9: 8372532 Brush_Engineered_Materials
[pem_pos=0, final_scores=0.5509798,
final_local_scores=0.5083237, pem=0.0, similarity=19.259108]
```

Sample: b'491_Salomon'

Salomon cuts refiner Q3 EPS view on margin concern . [1 1 NEW YORK 1] 1996-08-26 [2 2 Salomon Brothers 2] analyst Paul Ting said he cut his share earnings estimates on refiners for the third quarter on the belief the companies will face sharp revisions in refining and marketing margins . He cut his third-quarter share earnings estimate on : -- Diamond Shamrock Inc to \$ 0.38 from \$ 0.73 versus the Street 's consensus \$ 0.63 -- Sun Co to \$ 0.15 from \$ 0.85 versus the consensus \$ 0.63 -- Tosco Corp to \$ 0.95 from \$ 1.03 versus the consensus \$ 0.94 -- Total Petroleum (North America) Ltd to \$ 0.15 from \$ 0.46 versus the consensus \$ 0.33 -- And , [3 3 Valero Energy Corp 3] to \$ 0.27 from \$ 0.55 compared with the consensus \$ 0.40 .

```
1: 645042 New_York_City
2: 1043102 Salomon_Brothers
3: 306926 Valero_Energy_Corporation
1: 645042 New_York_City
[pem_pos=1, final_scores=0.37139267,
final_local_scores=0.4364463, pem=-1.7906687,
similarity=26.852076]
2: 1043102 Salomon_Brothers
[pem_pos=0, final_scores=0.44710124,
final_local_scores=0.53601825, pem=0.0, similarity=20.085567]
3: 306926 Valero_Energy_Corporation
[pem_pos=0, final_scores=0.39091718,
final_local_scores=0.4868579, pem=0.0, similarity=18.618523]
```

Sample: b'492_PRESALE'

PRESALE - Marion County Board of Education , W. Va . AMT :
3,250,000 DATE : 09/04/96 NYC Time : 1200 CUSIP : 569399
ISSUER : Marion County Board of Education ST : WV ISSUE :
Public School , Series 1996 TAX STAT:Exempt-ULT M / SP / F :
NA / NA / NA BOOK ENTRY : Y ENHANCEMENTS : None BANK QUAL : Y

DTD : 09/01/96 SURE BID : N DUE : 5/1/98-02 SR MGR : 1ST
CPN : 05/01/97 CALL : Non-Callable NIC DELIVERY : 9/17/96
approx ORDERS : PAYING AGENT : WesBanco Bank [1 1 Fairmont 1]
, [2 2 Fairmont 2] L.O. : [3 3 Steptoe & Johnson 3] , [4 4 Cl
arksburg 4] F.A. : Ferris , Baker Watts , Inc. , [5 5 Charles
ton 5] LAST SALE : \$ 7,330,000 (MBIA) 3/1/90 @ 6.14900 %
NIC ; 4yrs 4mos Avg ; BBI-7.27 % Year Amount Coupon Yield
Price Conc . 1998 575,000 1999 610,000 2000 650,000 2001
685,000 2002 730,000 COMPETITIVE PRE-SALE CONTRIBUTED BY J.J.
KENNY K-SHEETS :

1: 138476 Fairmont,_West_Virginia
2: 138476 Fairmont,_West_Virginia
3: 16350825 Steptoe_&_Johnson
4: 138420 Clarksburg,_West_Virginia
5: 57861 Charleston,_West_Virginia
1: 138476 Fairmont,_West_Virginia
[pem_pos=0, final_scores=0.6454488,
final_local_scores=0.62178075, pem=-0.95215344,
similarity=27.82285]
2: 138476 Fairmont,_West_Virginia
[pem_pos=0, final_scores=0.6552852,
final_local_scores=0.6391678, pem=-0.95215344,
similarity=28.341713]
3: 16350825 Steptoe_&_Johnson
[pem_pos=0, final_scores=0.35493362,
final_local_scores=0.44088328, pem=0.0, similarity=17.24655]
4: 138420 Clarksburg,_West_Virginia
[pem_pos=0, final_scores=0.4486845,
final_local_scores=0.42214638, pem=-1.2244449,
similarity=23.346125]
5: 57861 Charleston,_West_Virginia
[pem_pos=1, final_scores=0.25636595,
final_local_scores=0.21327251, pem=-2.012879,
similarity=21.400537]

Sample: b'493_BALANCE'
BALANCE - [1 1 Ohio 1] refunding bonds at \$ 290,000 . [2 2 ST
ATE OF OHIO 2] RE : \$ 70,375,000 (OHIO BUILDING AUTHORITY)
STATE FACILITIES REFUNDING BONDS 1996 SERIES A REPRICING OF
THE BALANCE OF THE BONDS IN THE ACCOUNT . 9,215,000.00 OCASEK
GOVERNMENT OFFICE BUILDING (C) MOODY 'S : A1 [3 3 S&P 3] :
AA- FITCH : AA- CONFIRMED CONFIRMED Delivery Date :
08/29/1996 Maturity Balance Coupon List 10/01 / 1998C 125M
4.50 4.20 6,045,000.00 VERN RIFFE CENTER (D) MOODY 'S : A1 [4 4 S&P 4] : AA- FITCH : AA- CONFIRMED CONFIRMED Delivery
Date : 08/29/1996 Maturity Balance Coupon List 10/01 / 1998D
165M 4.50 4.20 Grand Total : 290M [5 5 Goldman , Sachs & Co 5]
. A.G. Edwards & Sons , Inc . Banc One Capital Corporation
S.B.K- Brooks Investment Corp . Seasongood & Mayer -- U.S.
Municipal Desk , 212-859-1650
1: 22199 Ohio
2: 22199 Ohio

```
3: 164386 Standard_&Poor's
4: 164386 Standard_&Poor's
5: 335244 Goldman_Sachs
1: 22199 Ohio
    [pem_pos=0, final_scores=0.611255,
final_local_scores=0.6642082, pem=-0.19147073,
similarity=24.952257]
2: 22199 Ohio
    [pem_pos=0, final_scores=0.3670684,
final_local_scores=0.39747143, pem=0.0, similarity=15.951054]
3: 164386 Standard_&Poor's
    [pem_pos=0, final_scores=0.6069638,
final_local_scores=0.62590396, pem=0.0, similarity=22.767935]
4: 164386 Standard_&Poor's
    [pem_pos=0, final_scores=0.6434248,
final_local_scores=0.6657321, pem=0.0, similarity=23.956486]
5: 335244 Goldman_Sachs
    [pem_pos=0, final_scores=0.48755413,
final_local_scores=0.5331385, pem=0.0, similarity=19.99963]
```

Sample: b'494_PRESS'

```
PRESS DIGEST - [1 1 Washington Post 1] business - Aug 26 . [2
2 WASHINGTON 2] 1996-08-26 [3 3 The Washington Post 3]
carried only local business stories on August 26 , 1996 .
1: 102226 The_Washington_Post
2: 108956 Washington,_D.C.
3: 102226 The_Washington_Post
1: 102226 The_Washington_Post
    [pem_pos=0, final_scores=0.48282772,
final_local_scores=0.43302387, pem=0.0, similarity=17.012009]
2: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.5237251,
final_local_scores=0.57312745, pem=-1.7496481,
similarity=30.707838]
3: 102226 The_Washington_Post
    [pem_pos=0, final_scores=0.5056848,
final_local_scores=0.45799178, pem=0.0, similarity=17.757101]
```

Sample: b'495_Chelsea'

```
[1 1 Chelsea 1] makes political debut on [2 2 Clinton 2]
train trip . [3 3 CHILLICOTHE 3] , [4 4 Ohio 4] 1996-08-25 - [
5 5 Chelsea Clinton 5] , until now carefully shielded from
the exposure of public life , made her political debut on
Sunday on her father 's whistlestop train trip . [6 6 Chelsea
6] , 16 , was at President [7 7 Bill Clinton 7] 's side as he
rode the rails through parts of [8 8 West Virginia 8] , [9 9 K
entucky 9] and [10 10 Ohio 10] , and was introduced at every
stop . She even worked ropelines , shaking hands with excited
fans . [11 11 Hillary Rodham Clinton 11] saw her husband and
daughter off on the trip in [12 12 Huntington 12] , [13 13 We
st Virginia 13] and then went on to [14 14 Chicago 14] to
begin a rigorous [15 15 Democratic Convention 15] schedule .
```

Asked if [16 16 Chelsea 16] would have a prominent role in the campaign , [17 17 White House 17] spokesman [18 18 Mike Mc Curry 18] said : " She 'll do what she did today when she can . She has to go back to school . " The president 's daughter is going into her senior year of high school at [19 19 Sidwell Friends School 19] , a private school in [20 20 Washington 20] . McCurry said [21 21 Chelsea 21] has asked to go on the train trip and attend the convention where her father will be renominated , but said her exposure did not signal the start of a new political career . [22 22 Chelsea 22] " is a very poised young lady , but she 's not that much interested in politics , " the spokesman said .

1: 200683 Chelsea_Clinton
2: 3356 Bill_Clinton
3: 129782 Chillicothe,_Ohio
4: 22199 Ohio
5: 200683 Chelsea_Clinton
6: 200683 Chelsea_Clinton
7: 3356 Bill_Clinton
8: 32905 West_Virginia
9: 16846 Kentucky
10: 22199 Ohio
11: 5043192 Hillary_Rodham_Clinton
12: 93552 Huntington,_West_Virginia
13: 32905 West_Virginia
14: 6886 Chicago
15: 468102 Democratic_National_Convention
16: 200683 Chelsea_Clinton
17: 33057 White_House
18: 284483 Mike_McCurry_(press_secretary)
19: 925473 Sidwell_Friends_School
20: 108956 Washington,_D.C.
21: 200683 Chelsea_Clinton
22: 200683 Chelsea_Clinton
1: 200683 Chelsea_Clinton
 [pem_pos=0, final_scores=1.31418,
final_local_scores=1.3536248, pem=0.0, similarity=44.484577]
2: 3356 Bill_Clinton
 [pem_pos=1, final_scores=0.755194,
final_local_scores=0.77250034, pem=0.0, similarity=27.142664]
3: 129782 Chillicothe,_Ohio
 [pem_pos=0, final_scores=0.37565017,
final_local_scores=0.40672296, pem=-0.66169035,
similarity=19.825512]
4: 22199 Ohio
 [pem_pos=0, final_scores=0.6819398,
final_local_scores=0.70877004, pem=-0.19147073,
similarity=26.282072]
5: 200683 Chelsea_Clinton
 [pem_pos=0, final_scores=0.453557,
final_local_scores=0.40981835, pem=0.0, similarity=16.31951]
6: 200683 Chelsea_Clinton

```
[pem_pos=0, final_scores=1.4577312,
final_local_scores=1.5104327, pem=0.0, similarity=49.16404]
7: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.5847475,
final_local_scores=0.5804842, pem=0.0, similarity=21.41252]
8: 32905 West_Virginia
    [pem_pos=0, final_scores=0.49038053,
final_local_scores=0.5116918, pem=-0.30025622,
similarity=20.992456]
9: 16846 Kentucky
    [pem_pos=0, final_scores=0.3756042,
final_local_scores=0.39580363, pem=-0.28237545,
similarity=17.436886]
10: 22199 Ohio
    [pem_pos=0, final_scores=0.5806591,
final_local_scores=0.59813625, pem=-0.19147073,
similarity=22.98054]
11: 5043192 Hillary_Rodham_Clinton
    [pem_pos=0, final_scores=0.46199626,
final_local_scores=0.4521274, pem=0.0, similarity=17.582096]
12: 93552 Huntington,_West_Virginia
    [pem_pos=0, final_scores=0.20780025,
final_local_scores=0.20793062, pem=-1.8056281,
similarity=20.114063]
13: 32905 West_Virginia
    [pem_pos=0, final_scores=0.5095703,
final_local_scores=0.5326537, pem=-0.30025622,
similarity=21.618002]
14: 6886 Chicago
    [pem_pos=0, final_scores=0.64621025,
final_local_scores=0.66103846, pem=-0.20658566,
similarity=24.939865]
15: 468102 Democratic_National_Convention
    [pem_pos=0, final_scores=0.39307112,
final_local_scores=0.3887115, pem=-0.23190276,
similarity=16.950762]
16: 200683 Chelsea_Clinton
    [pem_pos=0, final_scores=1.3881788,
final_local_scores=1.4344573, pem=0.0, similarity=46.89678]
17: 33057 White_House
    [pem_pos=0, final_scores=0.46310735,
final_local_scores=0.44277966, pem=-0.0056634876,
similarity=17.33394]
18: 284483 Mike_McCurry_(press_secretary)
    [pem_pos=0, final_scores=0.44874197,
final_local_scores=0.49061406, pem=0.0, similarity=18.730616]
19: 925473 Sidwell_Friends_School
    [pem_pos=0, final_scores=0.52003336,
final_local_scores=0.58391696, pem=0.0, similarity=21.51496]
20: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.5373667,
final_local_scores=0.52257526, pem=-1.7496481,
```

```
similarity=29.199259]
21: 200683 Chelsea_Clinton
    [pem_pos=0, final_scores=1.5139698,
final_local_scores=1.5718648, pem=0.0, similarity=50.99729]
22: 200683 Chelsea_Clinton
    [pem_pos=0, final_scores=1.4865217,
final_local_scores=1.541882, pem=0.0, similarity=50.10255]
```

Sample: b'496_Dutch'

[1 1 Dutch 1] closing share market report . [2 2 AMSTERDAM 2]
1996-08-26 [3 3 Dutch 3] shares drifted to a lower close on Monday , dragged down by weakness in the domestic bond market and turnover depressed by the UK bank holiday and lack of any significant [4 4 U.S. 4] economic data . The [5 5 AEX 5] index of leading shares closed 4.54 points easier at 556.19 , the day 's low . The [6 6 Dutch 6] market had been on the defensive all day but a softer start to [7 7 Wall Street 7] did little to boost sentiment in late trade , dealers said . " It was always going to be a tough day with participation so low . But the only thing really worrying the market was the bonds , and that dragged us lower , " one dealer said . Stocks were down across the board , with Dutch PTT topping the volume list and closing down 1.90 guilders at 58.70 . The post and telecoms firm posted an 8.5 percent rise in first half earnings on Friday , just below analysts ' expectations . [8 8 IHC Caland 8] reported first half results well under forecasts on Monday and the shares suffered as a result . They were trading unchanged just before the release of figures but closed 2.40 guilders down at 80.70 after it reported net profits of 34.9 million guilders against 36.6 million last year and estimates ranging from 37.5 to 47.2 million . IHC also forecast post tax earnings rising 21 percent for the full year . Banking group [9 9 ING 9] traded ex-dividend today and finished 0.60 guilders weaker at 52.90 as a result . But [10 10 Nutricia 10] shrugged off its ex-div tag to soar a further 4.10 guilders to 214.40 continuing its explosive rally sparked by the 51 percent jump in first half net profits last week , which set the market alight on Friday , sending the shares up 18.40 at 210.00 by the close . Engineering concern Stork started the day well as the shares attracted some follow-through interest to the announcement late on Friday that its Fokker Aviation unit had won a major order . But the rally was short-lived and Stork ended just 0.20 up at 51.00 guilders . -- Amsterdam Newsroom +31 20 504 5000

1: 21148 Netherlands
2: 844 Amsterdam
3: 21148 Netherlands
4: 3434750 United_States
5: 1371940 AEX_index
6: 21148 Netherlands
7: 21560 New_York_Stock_Exchange

```
8: 2875996 SBM_Offshore
9: 230965 ING_Group
10: 2186586 Numico
1: 21148 Netherlands
    [pem_pos=0, final_scores=0.5687046,
final_local_scores=0.5511133, pem=-0.7943452,
similarity=24.855806]
2: 844 Amsterdam
    [pem_pos=0, final_scores=0.5833121,
final_local_scores=0.57841265, pem=-0.0642803,
similarity=21.700268]
3: 21148 Netherlands
    [pem_pos=0, final_scores=0.6434607,
final_local_scores=0.6327729, pem=-0.7943452,
similarity=27.292692]
4: 3434750 United_States
    [pem_pos=0, final_scores=0.44062626,
final_local_scores=0.466945, pem=-0.14075686,
similarity=18.78974]
5: 1371940 AEX_index
    [pem_pos=0, final_scores=0.4256345,
final_local_scores=0.53242266, pem=0.0, similarity=19.978268]
6: 21148 Netherlands
    [pem_pos=0, final_scores=0.6582666,
final_local_scores=0.64894617, pem=-0.7943452,
similarity=27.775333]
7: 21560 New_York_Stock_Exchange
    [pem_pos=7, final_scores=0.45721626,
final_local_scores=0.5096741, pem=-2.2964394,
similarity=31.787798]
8: 2875996 SBM_Offshore
    [pem_pos=0, final_scores=0.3922733,
final_local_scores=0.52901816, pem=0.0, similarity=19.87667]
9: 230965 ING_Group
    [pem_pos=0, final_scores=0.49116457,
final_local_scores=0.5382171, pem=0.0, similarity=20.151186]
10: 2186586 Numico
    [pem_pos=0, final_scores=0.30960315,
final_local_scores=0.40407085, pem=0.0, similarity=16.147993]
```

Sample: b'497_Tapie'
[1 1 Tapie 1] to quit [2 2 French 2] assembly seat as film
opens . [3 3 PARIS 3] 1996-08-26 Former businessman and
soccer boss [4 4 Bernard Tapie 4] said that he would give up
his seat in the National Assembly by Wednesday , the day a
film by [5 5 Claude Lelouch 5] in which he stars opens in [6
6 France 6] . " I will no longer be deputy by the time the
film opens , " he said in a broadcast interview . " Just
about , " he told [7 7 Europe 1 7] radio when asked whether
he had sent his letter of resignation to Assembly speaker [8 8
Philippe Seguin 8] . A Seguin spokeswoman confirmed that no
letter or call had yet been received . [9 9 Tapie 9] , 53 ,

was resigning just ahead of expected government action to eject him from the Assembly following a finding by the **1 Supreme Court** that he was bankrupt and thus ineligible for public office for a five-year period . [10 10 Tapie 10] , the target of a blizzard of legal actions over his now-destroyed business empire and the [11 11 Marseille 11] soccer team he once ran , has a starring role in Lelouche 's " Homme , femmes : mode d'emploi " (Men , women : instructions for use) . He plays a power-hungry lawyer in the movie described as " a tender and cruel comedy " by Lelouche , who is making his 35th film . " I have paid too dearly for mixing two careers , " [12 12 Tapie 12] said . " In [13 13 France 13] , you cannot be a film artist and a national politician at the same time . That is why I will no longer be deputy by the time the film opens . " Justice Minister [14 14 Jacques Toubon 14] began last month the formal process of ejecting [15 15 Tapie 15] from the [16 16 French 16] parliament as well as stripping him of his seat in the [17 17 European 17] parliament . The [18 18 French 18] procedure was expected to be completed before October 2 , when the National Assembly is to reconvene after a summer break , but the [19 19 European 19] procedure was expected to take longer . [20 20 Tapie 20] 's lawyer has said he intends to appeal to the [21 21 European Court of Human Rights 21] in an effort to prevent or delay the loss of his [22 22 European 22] seat . But such an appeal , even if the court were to accept the case , which it is not obliged to do , would not suspend enforcement of the [23 23 French 23] judgement against him . [24 24 Tapie 24] faces a probable spell in prison after he loses his parliamentary immunity , since two appeal courts have confirmed jail sentences of eight and six months against him for tax fraud and rigging a soccer match . He is appealing in both cases to the **2 Supreme Court** .

1: 324853 Bernard_Tapie
2: 5843419 France
3: 22989 Paris
4: 324853 Bernard_Tapie
5: 925369 Claude_Lelouch
6: 5843419 France
7: 1840470 Europe_1
8: 1615074 Philippe_Séguin
9: 324853 Bernard_Tapie
10: 324853 Bernard_Tapie
11: 40888948 Marseille
12: 324853 Bernard_Tapie
13: 5843419 France
14: 239898 Jacques_Toubon
15: 324853 Bernard_Tapie
16: 5843419 France
17: 9239 Europe
18: 5843419 France
19: 9239 Europe

```
20: 324853 Bernard_Tapie
21: 42622 European_Court_of_Human_Rights
22: 9239 Europe
23: 5843419 France
24: 324853 Bernard_Tapie
1: 220021 Court_of_Cassation_(France)
    [winner: 31737 Supreme_Court_of_the_United_States,
pem_pos=0, final_scores=-0.23006478,
final_local_scores=-0.17781031, pem=-0.9940822,
similarity=4.189472]
        RECALL MISS: 220021 Court_of_Cassation_(France)
2: 220021 Court_of_Cassation_(France)
    [winner: 31737 Supreme_Court_of_the_United_States,
pem_pos=0, final_scores=-0.18130448,
final_local_scores=-0.124547094, pem=-0.9940822,
similarity=5.7789526]
        RECALL MISS: 220021 Court_of_Cassation_(France)
1: 324853 Bernard_Tapie
    [pem_pos=0, final_scores=0.5818985,
final_local_scores=0.58898157, pem=0.0, similarity=21.666098]
2: 5843419 France
    [pem_pos=0, final_scores=0.7963571,
final_local_scores=0.7951738, pem=-0.80223274,
similarity=32.181953]
3: 22989 Paris
    [pem_pos=0, final_scores=0.5756188,
final_local_scores=0.57698315, pem=-0.115301244,
similarity=21.935066]
4: 324853 Bernard_Tapie
    [pem_pos=0, final_scores=0.8009936,
final_local_scores=0.8283098, pem=0.0, similarity=28.808128]
5: 925369 Claude_Lelouch
    [pem_pos=0, final_scores=0.48688173,
final_local_scores=0.5793993, pem=0.0, similarity=21.380144]
6: 5843419 France
    [pem_pos=0, final_scores=0.59989506,
final_local_scores=0.5805688, pem=-0.16240834,
similarity=22.298244]
7: 1840470 Europe_1
    [pem_pos=0, final_scores=0.455849,
final_local_scores=0.51472497, pem=0.0, similarity=19.450132]
8: 1615074 Philippe_Séguin
    [pem_pos=0, final_scores=0.57722473,
final_local_scores=0.65740633, pem=0.0, similarity=23.708029]
9: 324853 Bernard_Tapie
    [pem_pos=0, final_scores=0.6200159,
final_local_scores=0.63061905, pem=0.0, similarity=22.908644]
10: 324853 Bernard_Tapie
    [pem_pos=0, final_scores=0.6042728,
final_local_scores=0.6134221, pem=0.0, similarity=22.395452]
11: 40888948 Marseille
    [pem_pos=0, final_scores=0.42186075,
```

```
final_local_scores=0.4080993, pem=-0.2839409,
similarity=17.812325]
12: 324853 Bernard_Tapie
    [pem_pos=0, final_scores=0.53760177,
final_local_scores=0.54059416, pem=0.0, similarity=20.22212]
13: 5843419 France
    [pem_pos=0, final_scores=0.5880644,
final_local_scores=0.56764555, pem=-0.16240834,
similarity=21.91259]
14: 239898 Jacques_Toubon
    [pem_pos=0, final_scores=0.5341033,
final_local_scores=0.6622372, pem=0.0, similarity=23.852192]
15: 324853 Bernard_Tapie
    [pem_pos=0, final_scores=0.6681866,
final_local_scores=0.6832382, pem=0.0, similarity=24.478905]
16: 5843419 France
    [pem_pos=0, final_scores=0.7520496,
final_local_scores=0.7467746, pem=-0.80223274,
similarity=30.737623]
17: 9239 Europe
    [pem_pos=0, final_scores=0.46240216,
final_local_scores=0.51434666, pem=-0.25557294,
similarity=20.828688]
18: 5843419 France
    [pem_pos=0, final_scores=0.6681959,
final_local_scores=0.6551772, pem=-0.80223274,
similarity=28.004173]
19: 9239 Europe
    [pem_pos=0, final_scores=0.34641457,
final_local_scores=0.38764787, pem=-0.25557294,
similarity=17.047743]
20: 324853 Bernard_Tapie
    [pem_pos=0, final_scores=0.6423083,
final_local_scores=0.6549701, pem=0.0, similarity=23.635328]
21: 42622 European_Court_of_Human_Rights
    [pem_pos=0, final_scores=0.36683434,
final_local_scores=0.4580469, pem=0.0, similarity=17.758747]
22: 9239 Europe
    [pem_pos=0, final_scores=0.40166605,
final_local_scores=0.44800168, pem=-0.25557294,
similarity=18.848822]
23: 5843419 France
    [pem_pos=0, final_scores=0.7759549,
final_local_scores=0.7728874, pem=-0.80223274,
similarity=31.51688]
24: 324853 Bernard_Tapie
    [pem_pos=0, final_scores=0.64941376,
final_local_scores=0.6627318, pem=0.0, similarity=23.866953]
```

Sample: b'498_Bank'

[1 1 Bank of France 1] drains 3.9 bln Ffr at tender . [2 2 PA
RIS 2] 1996-08-26 The [3 3 Bank of France 3] drained 3.9

billion francs at a securities repurchase tender held on Monday to allocate funds for injection into the money market on Tuesday . It accepted bids for 44.3 billion francs in new liquidity , 3.9 billion less than the 48.2 billion leaving the market on Tuesday when a previous pact expires . The new pact expires on September 3 . The [4 4 Bank of France 4] said it allocated 13.4 billion francs to bidders offering Treasury bills as collateral , satisfying 3.4 percent of such demand . It allotted a further 30.9 billion to bidders putting up private paper , satisfying 12.5 percent of this demand . -- Paris Newsroom +33 1 4221 5452

1: 749262 Banque_de_France
2: 22989 Paris
3: 749262 Banque_de_France
4: 749262 Banque_de_France
1: 749262 Banque_de_France
 [pem_pos=0, final_scores=0.6572196,
final_local_scores=0.6063497, pem=0.0, similarity=22.184397]
2: 22989 Paris
 [pem_pos=0, final_scores=0.4624775,
final_local_scores=0.52649355, pem=-0.115301244,
similarity=20.428356]
3: 749262 Banque_de_France
 [pem_pos=0, final_scores=0.64216644,
final_local_scores=0.58990645, pem=0.0, similarity=21.693699]
4: 749262 Banque_de_France
 [pem_pos=0, final_scores=0.7074935,
final_local_scores=0.66126627, pem=0.0, similarity=23.823217]

Sample: b'499_RABOBANK'

[1 1 RABOBANK 1] [RABN.CN] SEES H2 NET GROWTH UNDER 10 PCT . [2 2 AMSTERDAM 2] 1996-08-26 [3 3 Dutch 3] co-operative bank Rabobank Nederland BA 's net profit growth might slow to less than than 10 year percent in the second half of 1996 , executive board chairman [4 4 Herman Wijffels 4] said on Monday . The unlisted bank earlier announced a 1996 interim net profit of 853 million guilders , up 21.5 percent on the 702 million guilders reported in the first half of 1995 . He said second-half profit growth would depend on customer demand for loans , which was already easing , and the performance of financial markets , which were strong in the first half and boosted securities , trading and underwriting income . " We expect reasonable growth in the second half . Maybe it will be in single-digits , " [5 5 Wijffels 5] told [6 6 Reuters 6] . He said growth in customer demand in the first half was 50 percent above normal but it was hard to maintain this pace . [7 7 Rabobank 7] 's net profit was 1.43 billion guilders in 1995 . The bank earlier warned that profit growth would slow in the second half , citing the increasing momentum in profit growth in the comparative period of 1995 and increased investment . But [8 8 Wijffels 8] was unable to quantify second half investment to improve and extend

```
domestic and offshore services . -- Garry West , [9 9 Amsterdam  
am 9] newsroom +31 20 504 5000  
1: 465956 Rabobank  
2: 844 Amsterdam  
3: 21148 Netherlands  
4: 8526629 Herman_Wijffels  
5: 8526629 Herman_Wijffels  
6: 18998750 Reuters  
7: 465956 Rabobank  
8: 8526629 Herman_Wijffels  
9: 844 Amsterdam  
1: 465956 Rabobank  
    [pem_pos=1, final_scores=0.3493706,  
final_local_scores=0.40369666, pem=-0.100605875,  
similarity=16.683937]  
2: 844 Amsterdam  
    [pem_pos=0, final_scores=0.5376652,  
final_local_scores=0.5919785, pem=-0.0642803,  
similarity=22.105099]  
3: 21148 Netherlands  
    [pem_pos=0, final_scores=0.4775579,  
final_local_scores=0.5318182, pem=-0.7943452,  
similarity=24.280003]  
4: 8526629 Herman_Wijffels  
    [pem_pos=0, final_scores=0.6246195,  
final_local_scores=0.6358742, pem=0.0, similarity=23.065468]  
5: 8526629 Herman_Wijffels  
    [pem_pos=0, final_scores=0.60111916,  
final_local_scores=0.6331871, pem=0.0, similarity=22.98528]  
6: 18998750 Reuters  
    [pem_pos=0, final_scores=0.2860908,  
final_local_scores=0.38800925, pem=0.0, similarity=15.668682]  
7: 465956 Rabobank  
    [pem_pos=1, final_scores=0.40494603,  
final_local_scores=0.4644044, pem=-0.100605875,  
similarity=18.495577]  
8: 8526629 Herman_Wijffels  
    [pem_pos=0, final_scores=0.5204472,  
final_local_scores=0.5450652, pem=0.0, similarity=20.355547]  
9: 844 Amsterdam  
    [pem_pos=0, final_scores=0.4964078,  
final_local_scores=0.54691106, pem=-0.0642803,  
similarity=20.760197]
```

Sample: b'500_PRESS'
PRESS DIGEST - [1 1 FRANCE 1] - [2 2 LE MONDE 2] AUG 26 . [3
3 PARIS 3] 1996-08-26 These are leading stories in Monday 's
afternoon daily [4 4 Le Monde 4] , dated Aug 27 . FRONT PAGE
-- [5 5 Ipsos 5] poll reports majority of [6 6 French 6]
public opinion sympathises with plight of [7 7 Africans 7]
seeking to renew or obtain work and residence permits ,
calling government " stubborn , " " confused " and " cold-

hearted . " BUSINESS PAGES -- [8 8 SNCF 8] railway trade unions want renegotiation of government bailout package , as [9 9 European Union 9] prepares more proposals to increase competition . -- World steel market shows signs of upturn . -- Paris Newsroom +33 1 42 21 53 81

1: 5843419 France
2: 127071 Le_Monde
3: 22989 Paris
4: 127071 Le_Monde
5: 3141209 Ipsos
6: 5843419 France
7: 5334607 Africa
8: 87110 SNCF
9: 9317 European_Union
1: 5843419 France
[pem_pos=0, final_scores=0.55082184,
final_local_scores=0.52646875, pem=-0.16240834,
similarity=20.683792]
2: 127071 Le_Monde
[pem_pos=0, final_scores=0.3758694,
final_local_scores=0.38462895, pem=0.0, similarity=15.567808]
3: 22989 Paris
[pem_pos=0, final_scores=0.5638159,
final_local_scores=0.56025326, pem=-0.115301244,
similarity=21.435814]
4: 127071 Le_Monde
[pem_pos=0, final_scores=0.4820503,
final_local_scores=0.5006155, pem=0.0, similarity=19.029078]
5: 3141209 Ipsos
[pem_pos=0, final_scores=0.5451661,
final_local_scores=0.65194154, pem=0.0, similarity=23.544949]
6: 5843419 France
[pem_pos=0, final_scores=0.7391033,
final_local_scores=0.7321376, pem=-0.80223274,
similarity=30.300825]
7: 5334607 Africa
[pem_pos=1, final_scores=0.6865147,
final_local_scores=0.75794226, pem=-1.1877377,
similarity=33.167324]
8: 87110 SNCF
[pem_pos=0, final_scores=0.28365034,
final_local_scores=0.3596249, pem=0.0, similarity=14.821637]
9: 9317 European_Union
[pem_pos=0, final_scores=0.4059767,
final_local_scores=0.44207263, pem=-0.005814796,
similarity=17.313663]

Sample: b'501_ATRIA'
ATRIA SEES H2 RESULT UP ON H1 . [1 1 HELSINKI 1] 1996-08-26 [
2 2 Finnish 2] foodstuffs group Atria Oy said in a statement
on Monday it expects its result to improve in the second half
of 1996 compared to the first half . " The result of the

second year-half is expected to improve on the early part of the year , " Atria said . Atria said earlier its January-June profit before extraordinary items , appropriations and taxes fell to 15 million markka from 39 in the first-half of 1995 .

1: 13696 Helsinki
2: 10577 Finland
1: 13696 Helsinki
[pem_pos=0, final_scores=0.67246455,
final_local_scores=0.65343714, pem=-0.015504186,
similarity=23.673895]
2: 10577 Finland
[pem_pos=0, final_scores=0.60495347,
final_local_scores=0.5796915, pem=-0.8112885,
similarity=25.800777]

Sample: b'502_BRIGHT-BELGIANS'

BRIGHT-BELGIANS SPEED AFTER [1 1 SCHUMACHER 1] 'S WIN . [2 2 BRUSSELS 2] [3 3 Michael Schumacher 3] 's victory in the [4 4 Belgian 4] [5 5 Formula One Grand Prix 5] at [6 6 Spa-Francorchamps 6] sparked a speeding epidemic on [7 7 Belgian 7] roads after the race was over . [8 8 Belga 8] news agency reported that police checked more than 3,000 drivers and booked 222 for speeding on their way home after the race . Some were clocked doing 180 kilometres an hour (112 miles per hour) , [9 9 Belga 9] said . [10 10 Schumacher 10] won the race in 1 hour 28 minutes 15.125 seconds at an average speed of 208.442 km / hour (130 m.p.h.) .

1: 20396 Michael_Schumacher
2: 3708 Brussels
3: 20396 Michael_Schumacher
4: 3343 Belgium
5: 10854 Formula_One
6: 948477 Circuit_de_Spa-Francorchamps
7: 3343 Belgium
8: 18360207 Belga_(news_agency)
9: 18360207 Belga_(news_agency)
10: 20396 Michael_Schumacher
1: 20396 Michael_Schumacher
[pem_pos=0, final_scores=0.64388114,
final_local_scores=0.63934255, pem=0.0, similarity=23.16897]
2: 3708 Brussels
[pem_pos=0, final_scores=0.33098733,
final_local_scores=0.3784647, pem=-0.025440974,
similarity=15.522207]
3: 20396 Michael_Schumacher
[pem_pos=0, final_scores=0.4982705,
final_local_scores=0.4671753, pem=0.0, similarity=18.031157]
4: 3343 Belgium
[pem_pos=0, final_scores=0.46565795,
final_local_scores=0.5150902, pem=-0.12403572,
similarity=20.135557]
5: 10854 Formula_One

```
[pem_pos=2, final_scores=0.24284665,
final_local_scores=0.2008236, pem=-1.271059,
similarity=16.994907]
6: 948477 Circuit_de_Spa-Francorchamps
    [pem_pos=0, final_scores=0.2930495,
final_local_scores=0.2741929, pem=0.0, similarity=12.272176]
7: 3343 Belgium
    [pem_pos=0, final_scores=0.5108462,
final_local_scores=0.5644515, pem=-0.12403572,
similarity=21.608597]
8: 18360207 Belga_(news_agency)
    [pem_pos=1, final_scores=0.37375534,
final_local_scores=0.49048972, pem=-0.17556477,
similarity=19.681652]
9: 18360207 Belga_(news_agency)
    [pem_pos=1, final_scores=0.42784345,
final_local_scores=0.5495727, pem=-0.17556477,
similarity=21.444807]
10: 20396 Michael_Schumacher
    [pem_pos=0, final_scores=0.54045683,
final_local_scores=0.5263672, pem=0.0, similarity=19.79756]
```

Sample: b'503_Thai'

[1 1 Thai 1] PM proposes Sept 18 for no-confidence debate . [2 2 BANGKOK 2] 1996-08-26 [3 3 Thai 3] Prime Minister [4 4 Banharn Silpa-archa 4] on Monday proposed September 18 as the date for parliamentary debate on an opposition no-confidence motion accusing him of incompetence . The president of parliament had earlier said September 11 could be set for the debate . The opposition motion against Banharn accuses him of being incompetent , lacking ethical leadership and alleges his administration is corrupt . His critics allege he may be attempting to delay the debate . Banharn has denied the accusations and said he is ready to clear himself in parliament . " In my opinion September 18 would be a convenient date for the government to answer questions . This has nothing to do with the accusation that I am trying to escape the debate , " Banharn told reporters after meeting coalition partners . Banharn 's 13-month-old , six-party coalition government controls 209 seats in the 391-seat lower house of parliament . Political infighting within Banharn 's [5 5 Chart Thai 5] party has raised doubts whether he can hold his supporters together and defeat the opposition motion , political analysts said . " We are still waiting to fix a date . September 18 is regarded as tentative because we still have not received the order to fix it on the agenda , " said an official at parliament 's agenda section . The last no-confidence debate against Banharn 's coalition in May was won by the government .

1: 30128 Thailand
2: 56615 Bangkok
3: 30128 Thailand

```
4: 7147085 Banharn_Silpa-archa
5: 1476804 Thai_Nation_Party
1: 30128 Thailand
    [pem_pos=2, final_scores=0.62760323,
final_local_scores=0.6230243, pem=-1.7504287,
similarity=32.201103]
2: 56615 Bangkok
    [pem_pos=0, final_scores=0.5173175,
final_local_scores=0.52414536, pem=-0.0005963434,
similarity=19.734499]
3: 30128 Thailand
    [pem_pos=2, final_scores=0.77062696,
final_local_scores=0.7792559, pem=-1.7504287,
similarity=36.863365]
4: 7147085 Banharn_Silpa-archa
    [pem_pos=0, final_scores=0.55865544,
final_local_scores=0.6447647, pem=0.0, similarity=23.330778]
5: 1476804 Thai_Nation_Party
    [pem_pos=0, final_scores=0.50244755,
final_local_scores=0.5889048, pem=0.0, similarity=21.663807]
```

Sample: b'504_Fontaine'

```
Fontaine - 6mth parent forecast . [1 1 TOKYO 1] 1996-08-26
Six months to August 31 , 1996 ( in billions of yen unless
specified ) LATEST PREVIOUS ACTUAL ( Parent ) FORECAST
FORECAST YEAR-AGO Sales 3.30 3.17 2.75 Current 400 million
260 million 231 million Net 170 million 170 million 142
million NOTE - Fontaine Co Ltd sells women " s fashion wigs .
1: 30057 Tokyo
1: 30057 Tokyo
    [pem_pos=0, final_scores=0.4163832,
final_local_scores=0.7274226, pem=-0.09283106,
similarity=26.302284]
```

Sample: b'505_Manila'

```
[1 1 Manila 1] international coconut oil prices . MANILA
1996-08-26 International Philippine coconut oil prices as
reported by the United Coconut Associations of the
Philippines ( dollars per tonne cif [1 1 Europe 1] ) . Buyers
Sellers Last Prev JulAug 775 787.50 unq unq AugSep 752.50
758.75 unq unq SepOct 733.75 743.50 unq unq OctNov unq 740
unq unq NovDec unq 732.50 unq unq
1: 9239 Europe
1: 72340 Metro_Manila
    [winner: 184334 Manila, pem_pos=0,
final_scores=0.09009366, final_local_scores=0.12760681,
pem=-0.11483788, similarity=8.522261]
    gt: 72340 Metro_Manila, pem_pos=1,
final_scores=-0.06088437, final_local_scores=0.0010989308,
pem=-2.832522, similarity=19.526201 ]
1: 9239 Europe
    [pem_pos=0, final_scores=0.5193142,
```

```
final_local_scores=0.60463405, pem=-0.086338006,  
similarity=22.602718]  
1: 184334 Manila  
    [pem_pos=0, final_scores=0.09009366,  
final_local_scores=0.12760681, pem=-0.11483788,  
similarity=8.522261]
```

Sample: b'506_RESEARCH'

RESEARCH ALERT - Aronkasei cut . [1 1 TOKYO 1] 1996-08-26
Nomura Research Institute Ltd downgraded Aronkasei Co Ltd to
a " 2 " rating from its previous " 1 " , market sources said
on Monday . In its three-grade rating system , the research
institute assigns a " 2 " rating to issues whose values it
sees moving within 10 percentage points in either direction
of the key 225-share [2 2 Nikkei 2] average over the next six
months . Nomura officials were not immediately available for
comment .

```
1: 30057 Tokyo  
2: 40293 Nikkei_225  
1: 30057 Tokyo  
    [pem_pos=0, final_scores=0.618259,  
final_local_scores=0.77953434, pem=-0.09283106,  
similarity=27.857403]  
2: 40293 Nikkei_225  
    [pem_pos=1, final_scores=0.33037,  
final_local_scores=0.46505928, pem=-0.46501607,  
similarity=20.49684]
```

Sample: b'507_PRESS'

PRESS DIGEST - [1 1 Tunisia 1] - Aug 26 . [2 2 TUNIS 2]
1996-08-26 These are the leading stories in the [3 3 Tunisian
3] press on Monday . [4 4 Reuters 4] has not verified these
stories and does not vouch for their accuracy . [5 5 LA PRESSE
5] - [6 6 English 6] langage to be taught as of the eighth
year of the primary school instead of the third year of the
secondary school . [7 7 LE TEMPS 7] - International Fair
opens in the northern city of [8 8 Beja 8] with the
participation of 16 foreign countries .

```
1: 30188 Tunisia  
2: 57659 Tunis  
3: 30188 Tunisia  
4: 18998750 Reuters  
5: 5141314 La_Presse_(French_newspaper)  
6: 8569916 English_language  
7: 1090060 Le_Temps  
8: 1403425 Béja  
1: 30188 Tunisia  
    [pem_pos=0, final_scores=0.47905123,  
final_local_scores=0.48379654, pem=-0.39291763,  
similarity=20.663914]  
2: 57659 Tunis  
    [pem_pos=0, final_scores=0.39771253,
```

```
final_local_scores=0.38930768, pem=-0.032719024,
similarity=15.885361]
3: 30188 Tunisia
    [pem_pos=0, final_scores=0.5724723,
final_local_scores=0.5858449, pem=-0.13068858,
similarity=22.283197]
4: 18998750 Reuters
    [pem_pos=0, final_scores=0.32341892,
final_local_scores=0.34368747, pem=0.0, similarity=14.346033]
5: 5141314 La_Presse_(French_newspaper)
    [pem_pos=1, final_scores=0.2658744,
final_local_scores=0.34622246, pem=-0.042628117,
similarity=14.6535]
6: 8569916 English_language
    [pem_pos=0, final_scores=0.46878654,
final_local_scores=0.5279636, pem=-0.640266,
similarity=23.327065]
7: 1090060 Le_Temps
    [pem_pos=0, final_scores=0.2868312,
final_local_scores=0.3666218, pem=0.0, similarity=15.030438]
8: 1403425 Béja
    [pem_pos=5, final_scores=0.24984606,
final_local_scores=0.32697135, pem=-2.142226,
similarity=25.496948]
```

Sample: b'508_U.N.'

[1 1 U.N. 1] official Ekeus heads for [2 2 Baghdad 2] . [3 3 MANAMA 3] 1996-08-26 Senior [4 4 United Nations 4] arms official [5 5 Rolf Ekeus 5] left [6 6 Bahrain 6] for [7 7 Bag hdad 7] on Monday for talks with [8 8 Iraqi 8] officials , a [9 9 U.N. 9] spokesman said . The spokesman said Ekeus , chairman of the [10 10 United Nations Special Commission 10] ([11 11 UNSCOM 11]) , would spend two or three days in [12 1 2 Iraq 12] but declined to give further details . [13 13 U.N. 13] ofcials have said Ekeus would hold talks with [14 14 Ira qi 14] Deputy Prime Minister Tareq Aziz and other officials as part of an agreement [15 15 Iraq 15] reached with the [16 1 6 United Nations 16] in June to hold higher level political talks with Ekeus . The [17 17 Security Council 17] on Friday asked [18 18 Iraq 18] to stop blocking arms inspectors search for concealed weapons or materials they believe were being shuttled around to avoid detection . Disarming [19 19 Iraq 19] of weapons of mass destruction under 1991 [20 20 Gulf War 20] ceasefire terms is a prerequisite before the lifting of crippling sanctions imposed on [21 21 Iraq 21] in 1990 for invading [22 22 Kuwait 22] .

1: 31769 United_Nations
2: 4492 Baghdad
3: 20481 Manama
4: 31769 United_Nations
5: 25837 Rolf_Ekéus
6: 18933277 Bahrain

7: 4492 Baghdad
8: 7515928 Iraq
9: 31769 United_Nations
10: 173504 United_Nations_Special_Commission
11: 173504 United_Nations_Special_Commission
12: 7515928 Iraq
13: 31769 United_Nations
14: 7515928 Iraq
15: 7515928 Iraq
16: 31769 United_Nations
17: 31956 United_Nations_Security_Council
18: 7515928 Iraq
19: 7515928 Iraq
20: 182000 Gulf_War
21: 7515928 Iraq
22: 7515890 Kuwait
1: 31769 United_Nations
[pem_pos=0, final_scores=0.47633737,
final_local_scores=0.44971675, pem=0.0, similarity=17.510159]
2: 4492 Baghdad
[pem_pos=0, final_scores=0.39911717,
final_local_scores=0.37043977, pem=0.0, similarity=15.144375]
3: 20481 Manama
[pem_pos=0, final_scores=0.42964035,
final_local_scores=0.4436965, pem=0.0, similarity=17.330502]
4: 31769 United_Nations
[pem_pos=0, final_scores=0.45989382,
final_local_scores=0.4317547, pem=0.0, similarity=16.974134]
5: 25837 Rolf_Ekéus
[pem_pos=0, final_scores=0.50794387,
final_local_scores=0.6034531, pem=0.0, similarity=22.097958]
6: 18933277 Bahrain
[pem_pos=0, final_scores=0.3944208,
final_local_scores=0.37075704, pem=-0.42385748,
similarity=17.458845]
7: 4492 Baghdad
[pem_pos=0, final_scores=0.34853822,
final_local_scores=0.31518996, pem=0.0, similarity=13.49561]
8: 7515928 Iraq
[pem_pos=0, final_scores=0.57956415,
final_local_scores=0.5299572, pem=-0.48469657,
similarity=22.540548]
9: 31769 United_Nations
[pem_pos=0, final_scores=0.5294285,
final_local_scores=0.50771075, pem=0.0, similarity=19.240814]
10: 173504 United_Nations_Special_Commission
[pem_pos=0, final_scores=0.43621814,
final_local_scores=0.45879716, pem=0.0, similarity=17.781136]
11: 173504 United_Nations_Special_Commission
[pem_pos=0, final_scores=0.40100175,
final_local_scores=0.42032862, pem=0.0, similarity=16.633158]
12: 7515928 Iraq

```
[pem_pos=0, final_scores=0.6172505,
final_local_scores=0.5711238, pem=-0.25379938,
similarity=22.513386]
13: 31769 United_Nations
    [pem_pos=0, final_scores=0.5454927,
final_local_scores=0.5252585, pem=0.0, similarity=19.764473]
14: 7515928 Iraq
    [pem_pos=0, final_scores=0.5792335,
final_local_scores=0.52959603, pem=-0.48469657,
similarity=22.52977]
15: 7515928 Iraq
    [pem_pos=0, final_scores=0.5926053,
final_local_scores=0.5442026, pem=-0.25379938,
similarity=21.710005]
16: 31769 United_Nations
    [pem_pos=0, final_scores=0.49269098,
final_local_scores=0.46758062, pem=0.0, similarity=18.043251]
17: 31956 United_Nations_Security_Council
    [pem_pos=0, final_scores=0.31001264,
final_local_scores=0.2969422, pem=0.0, similarity=12.95106]
18: 7515928 Iraq
    [pem_pos=0, final_scores=0.63288814,
final_local_scores=0.5882055, pem=-0.25379938,
similarity=23.02314]
19: 7515928 Iraq
    [pem_pos=0, final_scores=0.6585446,
final_local_scores=0.6162314, pem=-0.25379938,
similarity=23.859488]
20: 182000 Gulf_War
    [pem_pos=0, final_scores=0.3276934,
final_local_scores=0.29328865, pem=0.0, similarity=12.8420315]
21: 7515928 Iraq
    [pem_pos=0, final_scores=0.5374754,
final_local_scores=0.4839816, pem=-0.25379938,
similarity=19.91289]
22: 7515890 Kuwait
    [pem_pos=0, final_scores=0.4933135,
final_local_scores=0.45835185, pem=-0.3372397,
similarity=19.601809]
```

Sample: b'509_PRESS'
PRESS DIGEST - [1 1 Lebanon 1] - Aug 26 . [2 2 BEIRUT 2]
1996-08-26 These are the leading stories in the [3 3 Beirut 3]
press on Monday . [4 4 Reuters 4] has not verified these
stories and does not vouch for their accuracy . [5 5 AN-NAHAR
5] - The north [6 6 Lebanon 6] elections ... Almost no chance
at all for any complete ticket to win altogether and the
results will weaken some leaders . - The surprise ...
Opposition [7 7 Boutros Harb 7] scoring high in preliminary
results and former prime minister [8 8 Omar Karame 8] moves
backwards . - Fears of an [9 9 Israeli 9] operation causes
the redistribution of [10 10 Syrian 10] troops locations in [1

1 11 Lebanon 11] . [12 12 AS-SAFIR 12] - Parliament Speaker [1
3 13 Berri 13] : The occupied south should not be used as a
winning card in elections . - Prices of alimentary goods up
13.4 percent in 1996 . AL-ANWAR - Christian Maronite
Patriarch [14 14 Sfeir 14] : We fear a movement from
democracy to dictatorship . [15 15 AD-DIYAR 15] - A cabinet
minister : " Lebanon First " aims at splitting the Syrian-
Lebanese peace tracks with [16 16 Israel 16] . NIDA'A AL-
WATAN - Prime Minister [17 17 Hariri 17] : Elections are the
beginning of a long political life which we begin with an
incomplete ticket of 17 candidates . - The Lebanese
Association for the Democracy of Elections cited 51 incidents
of violation in the north [18 18 Lebanon 18] round .

1: 17771 Lebanon
2: 37428 Beirut
3: 37428 Beirut
4: 18998750 Reuters
5: 2180853 An-Nahar
6: 17771 Lebanon
7: 13608542 Boutros_Harb
8: 1090391 Omar_Karami
9: 9282173 Israël
10: 7515849 Syria
11: 17771 Lebanon
12: 4131343 As-Safir
13: 592775 Nabih_Berri
14: 1510400 Nasrallah_Boutros_Sfeir
15: 4273826 Ad-Diyar
16: 9282173 Israel
17: 394985 Rafic_Hariri
18: 17771 Lebanon
1: 17771 Lebanon
[pem_pos=0, final_scores=0.54778475,
final_local_scores=0.50044733, pem=-0.4097839,
similarity=21.252527]
2: 37428 Beirut
[pem_pos=0, final_scores=0.49572247,
final_local_scores=0.45786536, pem=-0.046614014,
similarity=18.006823]
3: 37428 Beirut
[pem_pos=0, final_scores=0.45319974,
final_local_scores=0.41141576, pem=-0.046614014,
similarity=16.620674]
4: 18998750 Reuters
[pem_pos=0, final_scores=0.31438094,
final_local_scores=0.3316039, pem=0.0, similarity=13.9854355]
5: 2180853 An-Nahar
[pem_pos=0, final_scores=0.637972,
final_local_scores=0.6878085, pem=0.0, similarity=24.615292]
6: 17771 Lebanon
[pem_pos=0, final_scores=0.5202711,
final_local_scores=0.47039288, pem=-0.4097839,

```
similarity=20.355642]
7: 13608542 Boutros_Harb
    [pem_pos=0, final_scores=0.38799232,
final_local_scores=0.4278956, pem=0.0, similarity=16.85897]
8: 1090391 Omar_Karami
    [pem_pos=0, final_scores=0.6398655,
final_local_scores=0.73108524, pem=0.0, similarity=25.906755]
9: 9282173 Israel
    [pem_pos=0, final_scores=0.5357316,
final_local_scores=0.524916, pem=-0.34790626,
similarity=21.646223]
10: 7515849 Syria
    [pem_pos=0, final_scores=0.5580267,
final_local_scores=0.54200405, pem=-0.60890055,
similarity=23.57549]
11: 17771 Lebanon
    [pem_pos=0, final_scores=0.5206813,
final_local_scores=0.470841, pem=-0.4097839,
similarity=20.369017]
12: 4131343 As-Safir
    [pem_pos=0, final_scores=0.59651595,
final_local_scores=0.63930994, pem=0.0, similarity=23.167997]
13: 592775 Nabih_Berri
    [pem_pos=2, final_scores=0.5112211,
final_local_scores=0.53702205, pem=-1.5646633,
similarity=28.624405]
14: 1510400 Nasrallah_Boutros_Sfeir
    [pem_pos=0, final_scores=0.3030526,
final_local_scores=0.32749274, pem=0.0, similarity=13.86275]
15: 4273826 Ad-Diyar
    [pem_pos=0, final_scores=0.64058393,
final_local_scores=0.7885315, pem=0.0, similarity=27.621065]
16: 9282173 Israel
    [pem_pos=0, final_scores=0.5445231,
final_local_scores=0.5345194, pem=-0.21279564,
similarity=21.198053]
17: 394985 Rafic_Hariri
    [pem_pos=0, final_scores=0.1416773,
final_local_scores=0.11010121, pem=-0.48811024,
similarity=10.029772]
18: 17771 Lebanon
    [pem_pos=0, final_scores=0.5174394,
final_local_scores=0.46729976, pem=-0.4097839,
similarity=20.263338]
```

Sample: b'510_PRESS'

PRESS DIGEST - [1 1 Malta 1] - Aug 26 . [2 2 VALLETTA 2]
1996-08-26 These are the leading stories in the [3 3 Maltese 3]
press on Monday . [4 4 Reuters 4] has not verified these
stories and does not vouch for their accuracy . [5 5 THE TIMES 5]
- Visitors slam bus and taxi drivers for cheating .
Tourists interviewed in [6 6 Malta 6] complain about over-

charging . [7 7 IN-NAZZJON 7] - Government considering measures for better road discipline . [8 8 Malta 8] , with a population of 365,000 , has 195,000 registered vehicles , with 80,000 new cars having been introduced on the congested roads in 10 years . - Five people arrested in [9 9 Romania 9] after drugs container found in [10 10 Malta 10] . The container , with 7.5 tonnes of cannabis , was found in [11 11 Malta Freeport 11] in transit from [12 12 Singapore 12] to [13 13 Romania 13] . [14 14 L-ORIZZONT 14] - Opposition leader [15 15 Alfred Sant 15] on steep rise in taxes over 10 years . He reiterates promise that a future [16 16 Labour 16] government will remove VAT .

1: 19137 Malta
2: 57664 Valletta
3: 19137 Malta
4: 18998750 Reuters
5: 39127 The_Times
6: 19137 Malta
7: 5694041 In-Nazzjon
8: 19137 Malta
9: 25445 Romania
10: 19137 Malta
11: 28223454 Malta_Freeport
12: 27318 Singapore
13: 25445 Romania
14: 5694096 L-Orizzont
15: 1174987 Alfred_Sant
16: 404834 Labour_Party_(Malta)
1: 19137 Malta
[pem_pos=0, final_scores=0.4689318,
final_local_scores=0.41343337, pem=-0.39222723,
similarity=18.560383]
2: 57664 Valletta
[pem_pos=0, final_scores=0.41674584,
final_local_scores=0.4120382, pem=-0.035661608,
similarity=16.579689]
3: 19137 Malta
[pem_pos=1, final_scores=0.41205364,
final_local_scores=0.35130265, pem=-1.3793737,
similarity=22.074533]
4: 18998750 Reuters
[pem_pos=0, final_scores=0.29960525,
final_local_scores=0.3418982, pem=0.0, similarity=14.292637]
5: 39127 The_Times
[pem_pos=0, final_scores=0.39834434,
final_local_scores=0.44791442, pem=0.0, similarity=17.456373]
6: 19137 Malta
[pem_pos=0, final_scores=0.4977947,
final_local_scores=0.44496173, pem=-0.39222723,
similarity=19.501251]
7: 5694041 In-Nazzjon
[pem_pos=0, final_scores=0.3615805,

```
final_local_scores=0.55110604, pem=0.0, similarity=20.535816]
8: 19137 Malta
    [pem_pos=0, final_scores=0.51827204,
final_local_scores=0.4673301, pem=-0.39222723,
similarity=20.168768]
9: 25445 Romania
    [pem_pos=0, final_scores=0.66498584,
final_local_scores=0.70724404, pem=-0.21241479,
similarity=26.35043]
10: 19137 Malta
    [pem_pos=0, final_scores=0.50071394,
final_local_scores=0.44815052, pem=-0.39222723,
similarity=19.59641]
11: 28223454 Malta_Freeport
    [pem_pos=0, final_scores=0.3535667,
final_local_scores=0.5008342, pem=0.0, similarity=19.035604]
12: 27318 Singapore
    [pem_pos=0, final_scores=0.45750433,
final_local_scores=0.4764676, pem=-0.1922004,
similarity=19.353672]
13: 25445 Romania
    [pem_pos=0, final_scores=0.7022502,
final_local_scores=0.7479497, pem=-0.21241479,
similarity=27.565168]
14: 5694096 L-Orizzont
    [pem_pos=0, final_scores=0.4146768,
final_local_scores=0.5928103, pem=0.0, similarity=21.780354]
15: 1174987 Alfred_Sant
    [pem_pos=0, final_scores=0.57028574,
final_local_scores=0.65790015, pem=0.0, similarity=23.722765]
16: 404834 Labour_Party_(Malta)
    [pem_pos=10, final_scores=0.09991993,
final_local_scores=0.09634918, pem=-3.4325604,
similarity=25.631763]
```

Sample: b'511_Tamils'

[1 1 Tamils 1] demonstrate outside [2 2 U.N. 2] headquarters . [3 3 GENEVA 3] 1996-08-26 Thousands of [4 4 Tamils 4] demonstrated outside the [5 5 United Nations 5] ' [6 6 European 6] headquarters in [7 7 Geneva 7] on Monday to appeal for [8 8 U.N. 8] recognition of their fight for independence from [9 9 Sri Lanka 9] . The demonstrators , said by police to number 6,000 , also urged the release of Nadarajah Muralidaran , Swiss-based leader of the the [10 10 Tamil Tiger 10] guerrillas , who has been held in a [11 11 Zurich 11] jail since April on charges of extortion . The demonstrators delivered an appeal to the [12 12 U.N. 12] human rights centre demanding an immediate end to " state terrorism " against [13 13 Tamils 13] and the [14 14 Liberation Tigers of Tamil Eelam 14] ([15 15 LTTE 15]) .

1: 300548 Tamil_people

2: 31769 United_Nations

```
3: 12521 Geneva
4: 300548 Tamil_people
5: 31769 United_Nations
6: 9239 Europe
7: 12521 Geneva
8: 31769 United_Nations
9: 26750 Sri_Lanka
10: 18606 Liberation_Tigers_of_Tamil_Eelam
11: 40334603 Zürich
12: 31769 United_Nations
13: 300548 Tamil_people
14: 18606 Liberation_Tigers_of_Tamil_Eelam
15: 18606 Liberation_Tigers_of_Tamil_Eelam
1: 300548 Tamil_people
    [pem_pos=0, final_scores=0.30497736,
final_local_scores=0.29636273, pem=-0.005057526,
similarity=12.961271]
2: 31769 United_Nations
    [pem_pos=0, final_scores=0.44340473,
final_local_scores=0.41865736, pem=0.0, similarity=16.583284]
3: 12521 Geneva
    [pem_pos=0, final_scores=0.48296112,
final_local_scores=0.48955536, pem=-0.24772356,
similarity=20.04618]
4: 300548 Tamil_people
    [pem_pos=0, final_scores=0.298303,
final_local_scores=0.289072, pem=-0.005057526,
similarity=12.743702]
5: 31769 United_Nations
    [pem_pos=0, final_scores=0.43103963,
final_local_scores=0.4051504, pem=0.0, similarity=16.18021]
6: 9239 Europe
    [pem_pos=0, final_scores=0.5025998,
final_local_scores=0.53623587, pem=-0.25557294,
similarity=21.481905]
7: 12521 Geneva
    [pem_pos=0, final_scores=0.55066174,
final_local_scores=0.5635081, pem=-0.24772356,
similarity=22.253078]
8: 31769 United_Nations
    [pem_pos=0, final_scores=0.54037434,
final_local_scores=0.52458197, pem=0.0, similarity=19.744286]
9: 26750 Sri_Lanka
    [pem_pos=0, final_scores=0.48788834,
final_local_scores=0.49315757, pem=-0.3027737,
similarity=20.453049]
10: 18606 Liberation_Tigers_of_Tamil_Eelam
    [pem_pos=0, final_scores=0.49787605,
final_local_scores=0.51514447, pem=0.0, similarity=19.462652]
11: 40334603 Zürich
    [pem_pos=0, final_scores=0.5937318,
final_local_scores=0.6684422, pem=-0.08839822,
```

```
similarity=24.518084]
12: 31769 United_Nations
    [pem_pos=0, final_scores=0.57203215,
final_local_scores=0.5591633, pem=0.0, similarity=20.77626]
13: 300548 Tamil_people
    [pem_pos=0, final_scores=0.35626417,
final_local_scores=0.35238582, pem=-0.005057526,
similarity=14.633112]
14: 18606 Liberation_Tigers_of_Tamil_Eelam
    [pem_pos=0, final_scores=0.5035404,
final_local_scores=0.5213319, pem=0.0, similarity=19.647297]
15: 18606 Liberation_Tigers_of_Tamil_Eelam
    [pem_pos=0, final_scores=0.46002752,
final_local_scores=0.47380072, pem=0.0, similarity=18.228872]
```

Sample: b'512_FOCUS'

FOCUS - Eurobourses end mixed but [1 1 London 1] recovers . Leonard Santorelli [2 2 LONDON 2] 1996-08-27 [3 3 European 3] bourses closed mixed on Tuesday with [4 4 London 4] clawing back most of the day 's losses despite an unsteady start on [5 5 wall Street 5] , hit by inflation worries . The dollar weakened during the day with many dealers sidelined because of uncertainty over [6 6 Tokyo 6] 's monetary direction ahead of the important [7 7 Japanese 7] Tankan economic survey out on Wednesday . But it recovered at the close of trade . Stocks in [8 8 London 8] started the week badly after a three-day weekend , slipping 0.3 percent , but bargain-hunters later moved in and the [9 9 FTSE 9] index recovered most of the lost ground to end only just in negative ground . Tuesday 's patchy showing in [10 10 London 10] followed a string of records last week , culminating in Friday 's trading high of 3,911 , fuelled by a wave of [11 11 European 11] interest rate cuts . The [12 12 London 12] bourse drew little help from the unsettled morning on [13 13 Wall Street 13] , which slipped in and out of positive ground after a stronger-than-expected August consumer confidence report refuelled inflation fears , pulling U.S. Treasuries back from their early peaks . Shortly after the report appeared showing the confidence index rising to 109.4 in August from a revised 107.0 in July , [14 14 Wall Street 14] relinquished virtually all its morning gains . " Treasuries remain very sensitive to any indication of a strong economy and we 're still in that summer doldrum period of light trading , " said Alan Ackerman , market strategist at Fahnestock & Co . " Consequently , stocks and bonds are both subject to rapid swings . " 1 Frankfurt was the one bright spot in [15 15 Euro pe 15] . Floor trading ended up 0.25 percent and the computerised IBIS index climbed nearly 0.4 percent , given a push by the performance of chemical shares . " The market is trading 100 percent on fundamentals at the moment as interest rate fantasy disappears -- it 's all company news , " said one trader . [16 16 French 16] shares ended slightly down

amid growing unease about a difficult autumn for the government which also weighed on the franc , dealers said . Bond prices were weaker and the franc was quoted at 3.4210 per mark for the first time since August 13 as worries about the government 's autumn budget and a weak [17 17 U.S. 17] currency lifted the mark and squeezed [18 18 French 18] investments . The dollar , which dropped sharply on Monday because of jitters over the [19 19 Japanese 19] Tankan survey , weakened further in quiet trading but regained losses to end the day close to Monday 's levels . " Besides the Tankan , there 's nothing really until the [20 20 U.S. 20] jobs numbers next week , " said a UK bank corporate dealer . " There is still the summer malaise hanging over the market . " Foreign exchange markets regard the Tankan report as an important indication of the country 's future monetary policy direction . If it points to weakness in the economy , the report will help the dollar regain some lost ground as speculation of a near-term rise in [21 21 Japanese 21] interest rates will evaporate . The [22 22 Japanese 22] discount rate is currently at a record low of 0.5 percent . The dollar was also pressured by the mark 's strength against [23 23 European 23] currencies amid renewed concerns over [24 24 Europe 24] 's economic and monetary union (EMU) timetable . CURRENCIES The dollar was at 1.4788 marks and 107.74 yen at the close of [25 25 European 25] trading compared with 1.4789 marks and 107.55 yen on Monday . STOCK MARKETS The Financial Times-Stock Exchange index of 100 leading [26 26 British 26] shares ended 1.8 points lower at 3,905.7 . In [27 27 Paris 27] , the CAC-40 share index finished down 2.43 at 2,017.99 . The 30-share [28 28 DAX 28] index in [29 29 Frankfurt 29] closed up 6.48 at 2,558.84 . PRECIOUS METALS Gold closed at \$ 388.55 an ounce , compared with Monday 's close of \$ 388.75 on international markets . Silver ended up one cent \$ 5.24 .

- 1: 17867 London
- 2: 17867 London
- 3: 9239 Europe
- 4: 17867 London
- 5: 21560 New_York_Stock_Exchange
- 6: 30057 Tokyo
- 7: 15573 Japan
- 8: 17867 London
- 9: 497717 FTSE_100_Index
- 10: 17867 London
- 11: 9239 Europe
- 12: 17867 London
- 13: 21560 New_York_Stock_Exchange
- 14: 21560 New_York_Stock_Exchange
- 15: 9239 Europe
- 16: 5843419 France
- 17: 3434750 United_States
- 18: 5843419 France

```
19: 15573 Japan
20: 3434750 United_States
21: 15573 Japan
22: 15573 Japan
23: 9239 Europe
24: 9239 Europe
25: 9239 Europe
26: 31717 United_Kingdom
27: 22989 Paris
28: 34153 DAX
29: 10992 Frankfurt
1: 34153 DAX
    [winner: 10992 Frankfurt, pem_pos=0,
final_scores=-0.17968413, final_local_scores=-0.23928429,
pem=-0.07062824, similarity=-2.6669202]
    RECALL MISS: 34153 DAX]
1: 17867 London
    [pem_pos=0, final_scores=0.66256326,
final_local_scores=0.64626527, pem=-0.09086737,
similarity=23.869707]
2: 17867 London
    [pem_pos=0, final_scores=0.77608854,
final_local_scores=0.7702744, pem=-0.09086737,
similarity=27.570387]
3: 9239 Europe
    [pem_pos=0, final_scores=0.43019682,
final_local_scores=0.39019626, pem=-0.25557294,
similarity=17.123793]
4: 17867 London
    [pem_pos=0, final_scores=0.4885915,
final_local_scores=0.45622754, pem=-0.09086737,
similarity=18.198603]
5: 21560 New_York_Stock_Exchange
    [pem_pos=7, final_scores=0.5834152,
final_local_scores=0.65041614, pem=-2.2964394,
similarity=35.987823]
6: 30057 Tokyo
    [pem_pos=0, final_scores=0.44720644,
final_local_scores=0.43682098, pem=-0.09283106,
similarity=17.630152]
7: 15573 Japan
    [pem_pos=0, final_scores=0.90229267,
final_local_scores=0.9131078, pem=-0.99076414,
similarity=36.7266]
8: 17867 London
    [pem_pos=0, final_scores=0.456721,
final_local_scores=0.4214139, pem=-0.09086737,
similarity=17.159695]
9: 497717 FTSE_100_Index
    [pem_pos=0, final_scores=0.352318,
final_local_scores=0.4270795, pem=0.0, similarity=16.834618]
10: 17867 London
```

```
[pem_pos=0, final_scores=0.5087599,
final_local_scores=0.47825855, pem=-0.09086737,
similarity=18.856052]
11: 9239 Europe
    [pem_pos=0, final_scores=0.34577578,
final_local_scores=0.29797912, pem=-0.25557294,
similarity=14.371849]
12: 17867 London
    [pem_pos=0, final_scores=0.43113118,
final_local_scores=0.39346087, pem=-0.09086737,
similarity=16.32552]
13: 21560 New_York_Stock_Exchange
    [pem_pos=7, final_scores=0.44474882,
final_local_scores=0.49894416, pem=-2.2964394,
similarity=31.467596]
14: 21560 New_York_Stock_Exchange
    [pem_pos=7, final_scores=0.40598243,
final_local_scores=0.4565978, pem=-2.2964394,
similarity=30.203897]
15: 9239 Europe
    [pem_pos=0, final_scores=0.63756585,
final_local_scores=0.61671543, pem=-0.086338006,
similarity=22.963251]
16: 5843419 France
    [pem_pos=0, final_scores=0.6889002,
final_local_scores=0.6889152, pem=-0.80223274,
similarity=29.010983]
17: 3434750 United_States
    [pem_pos=0, final_scores=0.495524,
final_local_scores=0.4633748, pem=-0.14075686,
similarity=18.683199]
18: 5843419 France
    [pem_pos=0, final_scores=0.6633312,
final_local_scores=0.66098505, pem=-0.80223274,
similarity=28.177492]
19: 15573 Japan
    [pem_pos=0, final_scores=0.86427134,
final_local_scores=0.8677059, pem=-0.99076414,
similarity=35.371716]
20: 3434750 United_States
    [pem_pos=0, final_scores=0.49690336,
final_local_scores=0.4648816, pem=-0.14075686,
similarity=18.728165]
21: 15573 Japan
    [pem_pos=0, final_scores=0.88998085,
final_local_scores=0.899659, pem=-0.99076414,
similarity=36.32526]
22: 15573 Japan
    [pem_pos=0, final_scores=0.7856541,
final_local_scores=0.7856979, pem=-0.99076414,
similarity=32.924435]
23: 9239 Europe
```

```
[pem_pos=0, final_scores=0.33872712,
final_local_scores=0.2902795, pem=-0.25557294,
similarity=14.142078]
24: 9239 Europe
    [pem_pos=0, final_scores=0.5549836,
final_local_scores=0.5265069, pem=-0.086338006,
similarity=20.27125]
25: 9239 Europe
    [pem_pos=0, final_scores=0.455459,
final_local_scores=0.41779137, pem=-0.25557294,
similarity=17.947285]
26: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5514347,
final_local_scores=0.51793545, pem=-0.48936236,
similarity=22.207169]
27: 22989 Paris
    [pem_pos=0, final_scores=0.6315505,
final_local_scores=0.65149575, pem=-0.115301244,
similarity=24.158672]
28: 34153 DAX
    [pem_pos=0, final_scores=0.6988807,
final_local_scores=0.8414614, pem=0.0, similarity=29.2006]
29: 10992 Frankfurt
    [pem_pos=0, final_scores=0.14856745,
final_local_scores=0.12512952, pem=-0.07062824,
similarity=8.207914]
```

Sample: b'513_More'

More hostages freed from hijacked [1 1 Sudanese 1] plane . [2 2 STANSTED 2] , [3 3 England 3] 1996-08-27 Armed hijackers believed to be [4 4 Iraqis 4] released between 60 and 70 people on Tuesday from a [5 5 Sudan Airways 5] plane carrying 199 passengers and crew that landed in [6 6 London 6] after being diverted on a flight from [7 7 Khartoum 7] to [8 8 Amman 8] , authorities said . A spokesman for [9 9 Stansted 9] airport said that the unidentified hijackers were demanding to see a [10 10 British-based 10] member of the Iraqi Community Association , called Mr Sadiki , and that police were trying to trace him . Police spokeswoman Kim White said police had already contacted Sadiki and were trying to arrange to bring him to [11 11 Stansted 11] , 30 miles (48 km) north-east of [12 12 London 12] , to talk to the hijackers . The airport spokesman said the six hijackers , who police said were armed with grenades and possibly other explosives , were believed to be [13 13 Iraqi 13] nationals . The hijackers started to release people for the Airbus plane in batches of 10 , starting with women and children , in what police described as a " controlled release " . Police said most of the passengers were [14 14 Sudanese 14] but that there were also an unknown number of [15 15 Iraqis 15] , Jordanians , [16 16 Palestinians 16] , [17 17 Syrians 17] and [18 18 Saudis 18] . Later they said the number of passengers

released from the plane had reached 80 .

1: 27421 Sudan

2: 65456 London_Stansted_Airport

3: 9316 England

4: 7515928 Iraq

5: 383994 Sudan_Airways

6: 17867 London

7: 17320 Khartoum

8: 45007 Amman

9: 65456 London_Stansted_Airport

10: 31717 United_Kingdom

11: 65456 London_Stansted_Airport

12: 17867 London

13: 7515928 Iraq

14: 27421 Sudan

15: 7515928 Iraq

16: 241405 State_of_Palestine

17: 7515849 Syria

18: 349303 Saudi_Arabia

1: 27421 Sudan

[pem_pos=0, final_scores=0.54830414,
final_local_scores=0.5265579, pem=-0.19893567,
similarity=20.885096]

2: 65456 London_Stansted_Airport

[pem_pos=0, final_scores=0.34189683,
final_local_scores=0.37110424, pem=-0.2670293,
similarity=16.61635]

3: 9316 England

[pem_pos=0, final_scores=0.6191365,
final_local_scores=0.64396006, pem=-0.19053856,
similarity=24.342943]

4: 7515928 Iraq

[pem_pos=1, final_scores=0.49191523,
final_local_scores=0.45842433, pem=-0.8129937,
similarity=22.191196]

5: 383994 Sudan_Airways

[pem_pos=0, final_scores=0.5247039,
final_local_scores=0.60537505, pem=0.0, similarity=22.155312]

6: 17867 London

[pem_pos=0, final_scores=0.45914596,
final_local_scores=0.4505121, pem=-0.09086737,
similarity=18.028044]

7: 17320 Khartoum

[pem_pos=0, final_scores=0.43439013,
final_local_scores=0.4039812, pem=-0.03966028,
similarity=16.360996]

8: 45007 Amman

[pem_pos=0, final_scores=0.43321198,
final_local_scores=0.406394, pem=0.0, similarity=16.217321]

9: 65456 London_Stansted_Airport

[pem_pos=0, final_scores=0.37802207,
final_local_scores=0.41056556, pem=-0.2670293,

```
similarity=17.793955]
10: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.59064215,
final_local_scores=0.582458, pem=0.0, similarity=21.471422]
11: 65456 London_Stansted_Airport
    [pem_pos=0, final_scores=0.35313314,
final_local_scores=0.3833782, pem=-0.2670293,
similarity=16.98263]
12: 17867 London
    [pem_pos=0, final_scores=0.49923205,
final_local_scores=0.49430007, pem=-0.09086737,
similarity=19.334763]
13: 7515928 Iraq
    [pem_pos=0, final_scores=0.5151735,
final_local_scores=0.48383045, pem=-0.48469657,
similarity=21.164034]
14: 27421 Sudan
    [pem_pos=0, final_scores=0.4402929,
final_local_scores=0.40857202, pem=-0.19893567,
similarity=17.36416]
15: 7515928 Iraq
    [pem_pos=1, final_scores=0.46169972,
final_local_scores=0.4254185, pem=-0.8129937,
similarity=21.206236]
16: 241405 State_of_Palestine
    [pem_pos=5, final_scores=0.33806628,
final_local_scores=0.35785472, pem=-1.7909594,
similarity=24.508327]
17: 7515849 Syria
    [pem_pos=0, final_scores=0.41459328,
final_local_scores=0.39379185, pem=-0.32786387,
similarity=17.624222]
18: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.34101874,
final_local_scores=0.2904069, pem=0.0, similarity=12.756035]
```

Sample: b'514_FEATURE'

FEATURE - " Eco terrorists " target UK builders . Edna Fernandes [1 1 LONDON 1] 1996-08-27 Ecological warfare has broken out across the [2 2 British 2] construction industry , striking some of the biggest corporates as activists give up peaceful protests and seek to hit builders where it hurts -- their profit margins . Described by one [3 3 British 3] company as " eco-terrorism " , it is seen as the new business risk of the 1990s . Famous names like Tarmac Plc , Costain Group Plc and [4 4 ARC 4] , a unit of conglomerate [5 5 Hanson Plc 5] , have all been targeted . Activist groups are no longer seen by [6 6 British 6] firms as a harmless , badly organised ragbag of students and hippies . " You only have to see them in action at protests , " said David Harding , spokesman at [7 7 ARC 7] , [8 8 Hanson 8] 's aggregates company . " They walk around with mobile phones and camera

equipment , they communicate and gather support for demos via the [9 9 Internet 9] -- we 're talking about a highly sophisticated organisation . " One road protestor under the codename Steady Eddie told construction journal " [10 10 Building 10] " earlier this year , " If it comes down to full-scale economic warfare , we will aim to drive them out of business . " As well as financial threats , companies also emphasise the " terror " tactics used . [11 11 Costain 11] 's contract to build the controversial [12 12 Newbury 12] bypass , which runs through a conservation area , has led to violent protests delaying building , bomb threats , staff intimidation and picketing of chief executive Alan Lovell 's home . A [13 13 Costain 13] spokesman told [14 14 Reuters 14] : " We 've had all sorts of protests at the head office and the chief executive 's house . But it 's when it gets to the (employee) families -- that it goes across the line . " Tactics used by some underground groups including the cryptic Berkshire Wood Elves , which distribute leaflets with instructions on home-made explosives , are now the subject of a police investigation . Other larger activist groups include Earth First , [15 15 The Land is Ours 15] , Alarm UK and Road Alert . The groups have targeted specific projects like the [1 6 16 Newbury 16] bypass and the M3 motorway through [17 17 Twyford Down 17] in the southern county of [18 18 Hampshire 18] . But they are also campaigning on broader issue such as stopping the government road building programme and out-of-town superstores which they say create more traffic , pollution and damage local communities . The government has slashed its road-building spending . Although protests may have contributed to the decision it has been seen primarily as economic rather than ecological . Graham Watts , chief executive of the [19 19 Construction Industry Council 19] , said : " I do n't think many firms involved in tendering for sensitive projects realise the impact environmental activity has on the cost of running a project . " But they are more alert than they were 3-4 years ago . There 's no doubt it 's a big issue now . " He says the damage comes in two forms : " Tangible -- in the form of extra costs , additional security , threats to staff and the more intangible damage caused by negative publicity . " Watts said the cost of protesting can be heavy once the company is locked into a contract . " I do often hear on the industry circuit of tales where the company tenders at low margins and the demonstrations which follow means they are running the project at a loss . " [20 20 ARC 20] says it 's not just contractors in the front line but also suppliers like itself . Its own quarries came under attack after it emerged that it may be a supplier for the [21 21 Newbury 21] bypass . " It was called the " First Battle of the Newbury bypass ' , " said [22 22 ARC 22] 's Harding . " We had 300 Earth First protestors invade and occupy our site . Hundreds of thousands of pounds (dollars) of damage was done in one day . Plus

there was the knock-on cost of lost production and extra security in future . " Simon Brown , analyst at investement bank [23 23 UBS 23] , said this new phenomenon has led to a change in the way the industry evaluates project risk . " When talking to [24 24 Tarmac 24] about the M3 link (through [25 25 Twyford Down 25]) they made it fairly clear that their risk assessment methods have been changed and now involve a very clear environmental risk analysis . " Harding says others have done the same . " As a result of eco-terrorism we are looking at controversial jobs more closely to see if the profit margins are wide enough to cover things like extra security . " For an industry already suffering from razor-thin margins , overcapacity and stagnant demand , eco-terrorism is the latest bizarre twist in the construction sector 's tale of woe .

```
1: 17867 London
2: 31717 United_Kingdom
3: 31717 United_Kingdom
4: 12152512 Amalgamated_Roadstone_Corporation
5: 1786946 Hanson_plc
6: 31717 United_Kingdom
7: 12152512 Amalgamated_Roadstone_Corporation
8: 1786946 Hanson_plc
9: 33139 World_Wide_Web
10: 9934632 Building_(magazine)
11: 3204747 Costain_Group
12: 125325 Newbury,_Berkshire
13: 3204747 Costain_Group
14: 18998750 Reuters
15: 18743234 The_Land_is_Ours
16: 125325 Newbury,_Berkshire
17: 277942 Twyford_Down
18: 13861 Hampshire
19: 9004141 Construction_Industry_Council
20: 12152512 Amalgamated_Roadstone_Corporation
21: 125325 Newbury,_Berkshire
22: 12152512 Amalgamated_Roadstone_Corporation
23: 25866628 UBS
24: 9058746 Tarmac_Group
25: 277942 Twyford_Down
1: 17867 London
    [pem_pos=0, final_scores=0.68103725,
final_local_scores=0.7121927, pem=-0.09086737,
similarity=25.837114]
2: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5060628,
final_local_scores=0.5102743, pem=-0.48936236,
similarity=21.978544]
3: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5492714,
final_local_scores=0.5574731, pem=-0.48936236,
similarity=23.387053]
```

```
4: 12152512 Amalgamated_Roadstone_Corporation
    [pem_pos=14, final_scores=0.8937269,
final_local_scores=0.9828821, pem=-3.4662347,
similarity=52.270798]
5: 1786946 Hanson_plc
    [pem_pos=0, final_scores=0.60008603,
final_local_scores=0.63391465, pem=0.0, similarity=23.00699]
6: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.58274233,
final_local_scores=0.594035, pem=-0.48936236,
similarity=24.478132]
7: 12152512 Amalgamated_Roadstone_Corporation
    [pem_pos=14, final_scores=0.897543,
final_local_scores=0.9870506, pem=-3.4662347,
similarity=52.395195]
8: 1786946 Hanson_plc
    [pem_pos=3, final_scores=0.697305,
final_local_scores=0.7401116, pem=-2.875713,
similarity=41.814693]
9: 33139 World_Wide_Web
    [pem_pos=1, final_scores=0.2799691,
final_local_scores=0.3560416, pem=-3.3911073,
similarity=33.156075]
10: 9934632 Building_(magazine)
    [pem_pos=1, final_scores=0.5853391,
final_local_scores=0.65991265, pem=-1.9268517,
similarity=34.26134]
11: 3204747 Costain_Group
    [pem_pos=0, final_scores=0.51358944,
final_local_scores=0.5444496, pem=0.0, similarity=20.337175]
12: 125325 Newbury,_Berkshire
    [pem_pos=0, final_scores=0.3214081,
final_local_scores=0.32380387, pem=-0.8724682,
similarity=18.497284]
13: 3204747 Costain_Group
    [pem_pos=0, final_scores=0.34081197,
final_local_scores=0.35571647, pem=0.0, similarity=14.705002]
14: 18998750 Reuters
    [pem_pos=0, final_scores=0.29520375,
final_local_scores=0.34944406, pem=0.0, similarity=14.51782]
15: 18743234 The_Land_is_Ours
    [pem_pos=0, final_scores=0.45298237,
final_local_scores=0.5738047, pem=0.0, similarity=21.213188]
16: 125325 Newbury,_Berkshire
    [pem_pos=0, final_scores=0.36351842,
final_local_scores=0.369803, pem=-0.8724682,
similarity=19.86999]
17: 277942 Twyford_Down
    [pem_pos=0, final_scores=0.5793285,
final_local_scores=0.6403048, pem=0.0, similarity=23.197685]
18: 13861 Hampshire
    [pem_pos=0, final_scores=0.5281523,
```

```
final_local_scores=0.5427395, pem=-0.3150438,
similarity=21.9994]
19: 9004141 Construction_Industry_Council
    [pem_pos=0, final_scores=0.46249056,
final_local_scores=0.57358694, pem=0.0, similarity=21.206692]
20: 12152512 Amalgamated_Roadstone_Corporation
    [pem_pos=14, final_scores=0.75841635,
final_local_scores=0.8350759, pem=-3.4662347,
similarity=47.859966]
21: 125325 Newbury,_Berkshire
    [pem_pos=0, final_scores=0.3838489,
final_local_scores=0.392011, pem=-0.8724682,
similarity=20.532719]
22: 12152512 Amalgamated_Roadstone_Corporation
    [pem_pos=14, final_scores=0.77842623,
final_local_scores=0.85693365, pem=-3.4662347,
similarity=48.512245]
23: 25866628 UBS
    [pem_pos=0, final_scores=0.29892445,
final_local_scores=0.3630157, pem=0.0, similarity=14.922826]
24: 9058746 Tarmac_Group
    [pem_pos=0, final_scores=0.44725907,
final_local_scores=0.47032034, pem=0.0, similarity=18.12501]
25: 277942 Twyford_Down
    [pem_pos=0, final_scores=0.5286398,
final_local_scores=0.5849351, pem=0.0, similarity=21.545343]
```

Sample: b'515_London'

London Carnival ends in high spirits . [1 1 LONDON 1]
1996-08-26 [2 2 London 2] 's annual [3 3 Notting_Hill_Carnival
3] , the largest in [4 4 Europe 4] and second in the world
only to [5 5 Rio 5] , ended peacefully on Monday with an
estimated 800,000 revellers singing and dancing the day away
in high spirits . Police said they made 30 arrests and there
were two stabbings . But there was no repeat of the ugly
scenes that used to scar the street festival , and police
praised the crowds over the two days of festivities as good-
natured . Around 400 police were wounded in riots in 1976
when the carnival , now in its 31st year , acquired its
darker reputation from which it is now only slowly
recovering . Shopkeepers still board up their windows and
many residents leave town for the weekend , but for four or
five years there has been no disorder and relatively little
crime .

```
1: 17867 London
2: 17867 London
3: 241103 Notting_Hill_Carnival
4: 9239 Europe
5: 25936 Rio_de_Janeiro
1: 17867 London
    [pem_pos=0, final_scores=0.7111349,
final_local_scores=0.6855673, pem=-0.09086737,
```

```
similarity=25.042559]
2: 17867 London
    [pem_pos=0, final_scores=0.5770841,
final_local_scores=0.5391372, pem=-0.09086737,
similarity=20.672792]
3: 241103 Notting_Hill_Carnival
    [pem_pos=0, final_scores=0.33280718,
final_local_scores=0.3872034, pem=0.0, similarity=15.644634]
4: 9239 Europe
    [pem_pos=0, final_scores=0.51076585,
final_local_scores=0.5496297, pem=-0.086338006,
similarity=20.961279]
5: 25936 Rio_de_Janeiro
    [pem_pos=0, final_scores=0.52511865,
final_local_scores=0.538239, pem=-1.4807417,
similarity=28.204342]
```

Sample: b'516_OFFICIAL'
OFFICIAL JOURNAL CONTENTS - OJ C 248 OF AUGUST 26 , 1996 . *
(Note - contents are displayed in reverse order to that in
the printed Journal) * ANNEX STATEMENT OF THE COUNCIL 'S
REASONS ANNEX I ANNEX A STATEMENT OF THE COUNCIL 'S REASONS
END OF DOCUMENT .

Sample: b'517_Indonesia'
[1 1 Indonesia 1] newspaper reports central bank scandal . [2
2 JAKARTA 2] 1996-08-27 [3 3 Indonesia 3] 's central bank
suffered seven billion rupiah (\$ 2.9 million) in losses
resulting from fake transactions , the [4 4 Jakarta Post 4]
reported on Tuesday . A [5 5 Bank Indonesia 5] spokeswoman
confirmed the newspaper report but declined to give further
details . The bank 's governor Sudradjat Djiwandono was
quoted as saying on Monday about 5.4 billion rupiah of the
seven billion rupiah had been recovered . The newspaper said
it was the central bank 's first public scandal in its 43-
year history . The paper said five people had been arrested
and police were looking for two more suspects . (\$ 1 = 2,341
rupiah)
1: 14579 Indonesia
2: 16275 Jakarta
3: 14579 Indonesia
4: 5717195 The_Jakarta_Post
5: 5044785 Bank_Indonesia
1: 14579 Indonesia
 [pem_pos=0, final_scores=0.57655066,
final_local_scores=0.5477696, pem=-0.23151328,
similarity=21.695257]
2: 16275 Jakarta
 [pem_pos=0, final_scores=0.47292137,
final_local_scores=0.45406985, pem=0.0, similarity=17.640062]
3: 14579 Indonesia

```
[pem_pos=0, final_scores=0.61336166,
final_local_scores=0.58798003, pem=-0.23151328,
similarity=22.895216]
4: 5717195 The_Jakarta_Post
    [pem_pos=0, final_scores=0.47167832,
final_local_scores=0.5135612, pem=0.0, similarity=19.415403]
5: 5044785 Bank_Indonesia
    [pem_pos=0, final_scores=0.48358864,
final_local_scores=0.5767723, pem=0.0, similarity=21.301748]
```

Sample: b'518_GOLF'

GOLF - PLAYERS DIVIDED ON CART REQUEST FOR OLAZABAL . [1 1 NO RTHAMPTON 1] , [2 2 England 2] 1996-08-27 [3 3 Seve Ballesteros 3] and [4 4 Colin Montgomerie 4] are divided on whether [5 5 Jose Maria Olazabal 5] should be allowed to return to the [6 6 European PGA Tour 6] using a motorised cart to transport him around the course . The [7 7 Spaniard 7] has not played for nearly a year because of rheumatoid arthritis in both his feet , and organisers of a pairs event to be staged in [8 8 Bordeaux 8] , [9 9 France 9] from October 17 to 20 have been asked to provide him with a buggy . " If the (Tour 's tournament) committee decides to change the rule I would not be against it , " said [10 10 Ballesteros 10] , Olazabal 's compatriot and [11 11 Ryder Cup 11] captain . But committee member [12 12 Montgomerie 12] said it could set an unhelpful precedent . " I know Olly 's situation is very unfortunate , but I do n't think we can start giving dispensations , " he said . " You 've got to have a rule for everybody and I do n't think it 's feasible . " The use of carts is generally prohibited in the professional game , and if Olazabal is allowed to use one in [13 13 Bordeaux 13] , he might then request one for the qualifying tournaments for next year 's [1 4 14 Ryder Cup 14] .

```
1: 63734 Northampton
2: 9316 England
3: 740993 Seve_Ballesteros
4: 379616 Colin_Montgomerie
5: 1519481 José_María_Olazábal
6: 841827 PGA_European_Tour
7: 26667 Spain
8: 4097 Bordeaux
9: 5843419 France
10: 740993 Seve_Ballesteros
11: 91198 Ryder_Cup
12: 379616 Colin_Montgomerie
13: 4097 Bordeaux
14: 91198 Ryder_Cup
1: 63734 Northampton
    [pem_pos=0, final_scores=0.41600227,
final_local_scores=0.47695738, pem=-0.5110059,
similarity=21.102003]
2: 9316 England
```

```
[pem_pos=0, final_scores=0.6096,
final_local_scores=0.6584813, pem=-0.19053856,
similarity=24.776287]
3: 740993 Seve_Ballesteros
    [pem_pos=0, final_scores=0.66896003,
final_local_scores=0.6590226, pem=0.0, similarity=23.756262]
4: 379616 Colin_Montgomerie
    [pem_pos=0, final_scores=0.7208033,
final_local_scores=0.72871107, pem=0.0, similarity=25.835905]
5: 1519481 José_María_Olazábal
    [pem_pos=0, final_scores=0.38001424,
final_local_scores=0.3871975, pem=0.0, similarity=15.644459]
6: 841827 PGA_European_Tour
    [pem_pos=0, final_scores=0.39211923,
final_local_scores=0.3857481, pem=0.0, similarity=15.601206]
7: 26667 Spain
    [pem_pos=1, final_scores=0.3195709,
final_local_scores=0.3516034, pem=-0.7552963,
similarity=18.689678]
8: 4097 Bordeaux
    [pem_pos=0, final_scores=0.5287292,
final_local_scores=0.5856348, pem=-0.40822297,
similarity=23.786203]
9: 5843419 France
    [pem_pos=0, final_scores=0.63444114,
final_local_scores=0.6920718, pem=-0.16240834,
similarity=25.625717]
10: 740993 Seve_Ballesteros
    [pem_pos=0, final_scores=0.6057841,
final_local_scores=0.5998544, pem=0.0, similarity=21.990564]
11: 91198 Ryder_Cup
    [pem_pos=0, final_scores=0.33329242,
final_local_scores=0.29658538, pem=0.0, similarity=12.9404125]
12: 379616 Colin_Montgomerie
    [pem_pos=0, final_scores=0.49173135,
final_local_scores=0.49030435, pem=0.0, similarity=18.721373]
13: 4097 Bordeaux
    [pem_pos=0, final_scores=0.50885034,
final_local_scores=0.56392014, pem=-0.40822297,
similarity=23.138195]
14: 91198 Ryder_Cup
    [pem_pos=0, final_scores=0.31095195,
final_local_scores=0.27218178, pem=0.0, similarity=12.21216]
```

Sample: b'519_SOCCER'

SOCCER - [1 1 POLAND 1] TIES [2 2 CYPRUS 2] 2-2 IN FRIENDLY MATCH . [3 3 BELCHATOW 3] , [4 4 Poland 4] 1996-08-27 [5 5 Poland 5] and [6 6 Cyprus 6] drew 2-2 (halftime 0-0) in a friendly soccer international on Tuesday . Scorers : [7 7 Poland 7] - [8 8 Krzysztof Warzycha 8] (46th minute) , [9 9 Mrcin Mieciel 9] (57th) [10 10 Cyprus 10] - Klimis Alexandrou (75th) , Kostas Malekos (80th) Attendance 3,000

```
1: 771860 Poland_national_football_team
2: 722717 Cyprus_national_football_team
3: 386721 Bełchatów
4: 22936 Poland
5: 771860 Poland_national_football_team
6: 722717 Cyprus_national_football_team
7: 771860 Poland_national_football_team
8: 3248501 Krzysztof_Warzycha
9: 8424753 Marcin_Mięciel
10: 722717 Cyprus_national_football_team
1: 771860 Poland_national_football_team
    [pem_pos=1, final_scores=0.6979191,
final_local_scores=0.67067665, pem=-3.319454,
similarity=42.15575]
2: 722717 Cyprus_national_football_team
    [pem_pos=1, final_scores=0.5260865,
final_local_scores=0.4998932, pem=-2.9483724,
similarity=35.04123]
3: 386721 Bełchatów
    [pem_pos=0, final_scores=0.2546665,
final_local_scores=0.4172169, pem=0.0, similarity=16.540297]
4: 22936 Poland
    [pem_pos=0, final_scores=0.51555544,
final_local_scores=0.5865306, pem=-0.15333635,
similarity=22.426823]
5: 771860 Poland_national_football_team
    [pem_pos=1, final_scores=0.5520245,
final_local_scores=0.51130897, pem=-3.319454,
similarity=37.3999]
6: 722717 Cyprus_national_football_team
    [pem_pos=1, final_scores=0.5832123,
final_local_scores=0.5622945, pem=-2.9483724,
similarity=36.903408]
7: 771860 Poland_national_football_team
    [pem_pos=1, final_scores=0.84606713,
final_local_scores=0.8325059, pem=-3.319454,
similarity=46.985058]
8: 3248501 Krzysztof_Warzycha
    [pem_pos=0, final_scores=0.44917047,
final_local_scores=0.5244925, pem=0.0, similarity=19.741615]
9: 8424753 Marcin_Mięciel
    [pem_pos=0, final_scores=0.7590918,
final_local_scores=0.8940217, pem=0.0, similarity=30.769102]
10: 722717 Cyprus_national_football_team
    [pem_pos=1, final_scores=0.64367753,
final_local_scores=0.62834346, pem=-2.9483724,
similarity=38.874443]
```

Sample: b'520_SOCCER'

SOCCER - THOMSON RESIGNS AS MANAGER OF [1 1 RAITH ROVERS 1] .
[2 2 KIRKCALDY 2] , [3 3 Scotland 3] 1996-08-27 [4 4 Jimmy Th
omson 4] became [1 1 Scotland 1] 's first managerial casualty

of the season on Tuesday when he quit [5 5 Raith Rovers 5] , bottom of the premier division . Thomson resigned after the club 's directors asked him to return to his previous position as youth team coach . He had been in charge for six months . [6 6 Raith 6] lost their first two games of the season away to [7 7 Rangers 7] and [8 8 Celtic 8] , then crashed 3-0 at home to [9 9 Motherwell 9] on Saturday . A club statement said : " The directors of Raith Rovers FC invited [10 10 Jimmy Thomson 10] to relinquish the post of manager and to resume his former position as youth team coach . " Regrettably Jimmy has felt unable to accept that offer , and has accordingly left the club . " Thomson said : " I am leaving with my dignity and my pride intact , and that is very important to me . " While not agreeing with the directors ' decision , I respect their reasons for making it . "

```
1: 880348 Raith_Rovers_F.C.
2: 355074 Kirkcaldy
3: 26994 Scotland
4: 23213228 Jimmy_Thomson_(footballer)
5: 880348 Raith_Rovers_F.C.
6: 880348 Raith_Rovers_F.C.
7: 69713 Rangers_F.C.
8: 69708 Celtic_F.C.
9: 215398 Motherwell_F.C.
10: 23213228 Jimmy_Thomson_(footballer)
1: 226426 Scottish_Premier_League
    [winner: 578101 Scotland_national_football_team,
pem_pos=1, final_scores=0.21410869,
final_local_scores=0.16220915, pem=-2.6956348,
similarity=23.58964]
    gt: 226426 Scottish_Premier_League, pem_pos=12,
final_scores=0.17045327, final_local_scores=0.10100961,
pem=-4.515878, similarity=31.662088 ]
1: 880348 Raith_Rovers_F.C.
    [pem_pos=0, final_scores=0.40649325,
final_local_scores=0.34193417, pem=0.0, similarity=14.293711]
2: 355074 Kirkcaldy
    [pem_pos=0, final_scores=0.54661745,
final_local_scores=0.5629312, pem=0.0, similarity=20.888702]
3: 26994 Scotland
    [pem_pos=0, final_scores=0.42316335,
final_local_scores=0.45620728, pem=-0.22905664,
similarity=18.949493]
4: 23213228 Jimmy_Thomson_(footballer)
    [pem_pos=2, final_scores=0.45992917,
final_local_scores=0.5263592, pem=-0.8737894,
similarity=24.549124]
5: 880348 Raith_Rovers_F.C.
    [pem_pos=0, final_scores=0.34192514,
final_local_scores=0.27140328, pem=0.0, similarity=12.188929]
6: 880348 Raith_Rovers_F.C.
```

```
[pem_pos=0, final_scores=0.44214773,
final_local_scores=0.38088125, pem=0.0, similarity=15.455969]
7: 69713 Rangers_F.C.
    [pem_pos=0, final_scores=0.63536096,
final_local_scores=0.60251135, pem=-0.30191645,
similarity=23.711721]
8: 69708 Celtic_F.C.
    [pem_pos=0, final_scores=0.60138273,
final_local_scores=0.57160026, pem=-0.9785444,
similarity=26.46888]
9: 215398 Motherwell_F.C.
    [pem_pos=0, final_scores=0.6634503,
final_local_scores=0.6262748, pem=-0.3361376,
similarity=24.606972]
10: 23213228 Jimmy_Thomson_(footballer)
    [pem_pos=2, final_scores=0.5159856,
final_local_scores=0.5875923, pem=-0.8737894,
similarity=26.376442]
1: 578101 Scotland_national_football_team
    [pem_pos=1, final_scores=0.21410869,
final_local_scores=0.16220915, pem=-2.6956348,
similarity=23.58964]
```

Sample: b'521_RUGBY'

```
[1 1 RUGBY UNION 1] - [2 2 GRIQUALAND WEST 2] AND [3 3 NEW ZEALAND 3] DRAW IN TOUR MATCH . [4 4 KIMBERLEY 4] , [5 5 South Africa 5] 1996-08-27 [6 6 Griqualand West 6] and [7 7 New Zealand 7] drew 18-18 ( halftime 6-10 ) in their rugby union tour match on Tuesday . Scorers : [8 8 Griqualand West 8] - Tries : Andre Cloete , Leon van der Wath . Conversion : Boeta Wessels . Penalties : Wessels 2 . [9 9 New Zealand 9] - Tries : Scott McLeod , [10 10 Glen Osborne 10] . Conversion : Jon Preston . Penalties : Preston 2 .
1: 25405 Rugby_union
2: 888197 Griqualand_West
3: 154957 New_Zealand_national_rugby_union_team
4: 313628 Kimberley,_Northern_Cape
5: 17416221 South_Africa
6: 888197 Griqualand_West
7: 154957 New_Zealand_national_rugby_union_team
8: 888197 Griqualand_West
9: 154957 New_Zealand_national_rugby_union_team
10: 11940740 Glen_Osborne
1: 25405 Rugby_union
    [pem_pos=0, final_scores=0.4521882,
final_local_scores=0.46911114, pem=0.0, similarity=18.088924]
2: 888197 Griqualand_West
    [pem_pos=0, final_scores=0.28553873,
final_local_scores=0.26481164, pem=-0.045276318,
similarity=12.23844]
3: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=4, final_scores=0.56699157,
```

```
final_local_scores=0.5485302, pem=-3.8192244,
similarity=41.228485]
4: 313628 Kimberley,_Northern_Cape
    [pem_pos=0, final_scores=0.40076727,
final_local_scores=0.4031027, pem=-0.77688056,
similarity=20.343899]
5: 17416221 South_Africa
    [pem_pos=0, final_scores=0.56877655,
final_local_scores=0.6037333, pem=-0.26251382,
similarity=23.53391]
6: 888197 Griqualand_West
    [pem_pos=0, final_scores=0.50848156,
final_local_scores=0.5083428, pem=-0.045276318,
similarity=19.505896]
7: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=4, final_scores=0.6268028,
final_local_scores=0.6138649, pem=-3.8192244,
similarity=43.1782]
8: 888197 Griqualand_West
    [pem_pos=0, final_scores=0.5696421,
final_local_scores=0.57515144, pem=-0.045276318,
similarity=21.4996]
9: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=4, final_scores=0.63137007,
final_local_scores=0.6188539, pem=-3.8192244,
similarity=43.327084]
10: 11940740 Glen_Osborne
    [pem_pos=1, final_scores=0.44716072,
final_local_scores=0.5631375, pem=-0.27698678,
similarity=22.401155]
```

Sample: b'522_SOCCER'

SOCCER - [1 1 BALL 1] RESIGNS AS [2 2 MANCHESTER CITY 2]
MANAGER . [3 3 MANCHESTER 3] , [4 4 England 4] 1996-08-27
Former [5 5 England 5] midfielder [6 6 Alan Ball 6] resigned
as manager of first division side [7 7 Manchester City 7] on
Monday night , the club said . [8 8 Ball 8] , appointed in
July 1995 , was unable to prevent [9 9 City 9] being
relegated from the premier league last season and his record
read 13 wins , 14 draws and 22 losses in 49 games . They have
lost two of their three matches so far this season . Club
secretary Bernard Halford said in a statement : " The
chairman and Board would like to place on record their
appreciation of his endeavours and efforts whilst in his
period of office and wish him well in the future . "
1: 899013 Alan_Ball,_Jr.
2: 165813 Manchester_City_F.C.
3: 20206 Manchester
4: 9316 England
5: 9904 England_national_football_team
6: 899013 Alan_Ball,_Jr.
7: 165813 Manchester_City_F.C.

```

8: 899013 Alan_Ball,_Jr.
9: 165813 Manchester_City_F.C.
1: 899013 Alan_Ball,_Jr.
    [pem_pos=1, final_scores=0.7323535,
final_local_scores=0.7093467, pem=-0.6195933,
similarity=28.627478]
2: 165813 Manchester_City_F.C.
    [pem_pos=0, final_scores=0.39727777,
final_local_scores=0.34432578, pem=0.0, similarity=14.365081]
3: 20206 Manchester
    [pem_pos=0, final_scores=0.52023906,
final_local_scores=0.5660528, pem=-0.3209542,
similarity=22.727257]
4: 9316 England
    [pem_pos=0, final_scores=0.503494,
final_local_scores=0.56692237, pem=-0.19053856,
similarity=22.043985]
5: 9904 England_national_football_team
    [pem_pos=1, final_scores=0.4373837,
final_local_scores=0.39468002, pem=-2.776464,
similarity=30.966595]
6: 899013 Alan_Ball,_Jr.
    [pem_pos=1, final_scores=0.59013814,
final_local_scores=0.5470471, pem=-0.6195933,
similarity=23.784134]
7: 165813 Manchester_City_F.C.
    [pem_pos=0, final_scores=0.3469212,
final_local_scores=0.28530744, pem=0.0, similarity=12.603856]
8: 899013 Alan_Ball,_Jr.
    [pem_pos=1, final_scores=1.0453427,
final_local_scores=1.05124, pem=-0.6195933,
similarity=38.830257]
9: 165813 Manchester_City_F.C.
    [pem_pos=12, final_scores=0.58280176,
final_local_scores=0.54297113, pem=-4.187771,
similarity=43.066803]

```

Sample: b'523_SOCCER'

SOCCER - [1 1 ENGLISH 1] LEAGUE STANDINGS . [2 2 LONDON 2]

1996-08-27 [3 3 English 3] league soccer standings after
Tuesday 's matches (tabulated under played , won , drawn ,
lost , goals for , against , points) : Division one [4 4 Tra
nmere 4] 3 2 1 0 6 3 7 [5 5 Bolton 5] 3 2 1 0 5 2 7 [6 6 Barn
sley 6] 2 2 0 0 5 2 6 [7 7 Wolverhampton 7] 2 2 0 0 4 1 6 [8 8
Queens Park Rangers 8] 2 2 0 0 4 2 6 [9 9 Stoke 9] 2 2 0 0 4
2 6 [10 10 Norwich 10] 3 2 0 1 4 3 6 [11 11 Ipswich 11] 3 1 1
1 6 4 4 [12 12 Birmingham 12] 2 1 1 0 5 4 4 [13 13 Crystal Pa
lace 13] 3 1 1 1 3 2 4 [14 14 Oxford 14] 3 1 0 2 6 3 3 1 Brad
ford 2 1 0 1 3 2 3 [15 15 Huddersfield 15] 2 1 0 1 3 3 3 [16 1
6 Portsmouth 16] 3 1 0 2 3 5 3 [17 17 Reading 17] 2 1 0 1 3 5
3 [18 18 Manchester City 18] 3 1 0 2 2 3 3 [19 19 West Bromwi
ch 19] 3 0 2 1 2 3 2 [20 20 Port Vale 20] 3 0 2 1 2 4 2 [21 21

Sheffield United 21] 2 0 1 1 4 5 1 [22 22 Grimsby 22] 3 0 1 2
 4 7 1 [23 23 Charlton 23] 2 0 1 1 1 3 1 [24 24 Swindon 24] 2
 0 1 1 1 3 1 [25 25 Southend 25] 3 0 1 2 1 7 1 [26 26 Oldham 26
] 2 0 0 2 2 5 0 Divisionn two 2 Plymouth 3 2 1 0 8 5 7 [27 27
 Brentford 27] 3 2 1 0 6 3 7 [28 28 Shrewsbury 28] 3 2 1 0 6 3
 7 [29 29 Bury 29] 3 2 1 0 4 2 7 [30 30 Burnley 30] 3 2 0 1 5
 5 6 [31 31 Bournemouth 31] 3 2 0 1 4 3 6 [32 32 Blackpool 32]
 3 2 0 1 3 2 6 [33 33 Chesterfield 33] 3 2 0 1 3 2 6 [34 34 Mi
 llwall 34] 3 1 1 1 5 4 4 [35 35 Crewe 35] 3 1 1 1 4 4 4 [36 36
 Gillingham 36] 3 1 1 1 4 5 4 [37 37 Preston 37] 3 1 1 1 3 3 4
 [38 38 Notts County 38] 2 1 1 0 2 1 4 [39 39 Bristol Rovers 39
] 2 1 1 0 1 0 4 [40 40 Bristol City 40] 3 1 0 2 7 4 3 3 York
 3 1 0 2 5 6 3 [41 41 Watford 41] 3 1 0 2 2 5 3 [42 42 Wrexham
 42] 2 0 2 0 5 5 2 [43 43 Wycombe 43] 3 0 2 1 2 3 2 [44 44 Rot
 herham 44] 3 0 1 2 3 5 1 [45 45 Peterborough 45] 2 0 1 1 2 3
 1 [46 46 Walsall 46] 3 0 1 2 2 4 1 [47 47 Stockport 47] 3 0 1
 2 0 2 1 [48 48 Luton 48] 3 0 0 3 3 10 0 Division three [49 49
 Hartlepool 49] 3 2 1 0 6 3 7 [50 50 Wigan 50] 3 2 1 0 5 2 7 [5
 1 51 Hull 51] 3 2 1 0 4 2 7 [52 52 Carlisle 52] 3 2 1 0 2 0 7
 [53 53 Fulham 53] 3 2 0 1 4 3 6 [54 54 Scunthorpe 54] 3 2 0 1
 2 2 6 [55 55 Scarborough 55] 3 1 2 0 4 2 5 4 Exeter 3 1 2 0 4
 3 5 5 Cambridge 3 1 2 0 3 2 5 [56 56 Darlington 56] 3 1 1 1 7
 5 4 [57 57 Northampton 57] 3 1 1 1 5 3 4 [58 58 Barnet 58] 3
 1 1 1 4 2 4 [59 59 Chester 59] 3 1 1 1 4 3 4 [60 60 Torquay 60
] 3 1 1 1 3 3 4 [61 61 Cardiff 61] 3 1 1 1 1 2 4 [62 62 Swans
 ea 62] 3 1 0 2 3 7 3 [63 63 Brighton 63] 3 1 0 2 2 5 3 [64 64
 Hereford 64] 3 1 0 2 1 2 3 6 Lincoln 3 0 2 1 3 4 2 [65 65 Col
 chester 65] 3 0 2 1 1 3 2 [66 66 Rochdale 66] 3 0 1 2 2 4 1 [6
 7 67 Mansfield 67] 3 0 1 2 2 6 1 [68 68 Doncaster 68] 3 0 1 2
 1 3 1 [69 69 Leyton Orient 69] 3 0 1 2 1 3 1

- 1: 9316 England
- 2: 17867 London
- 3: 9316 England
- 4: 69219 Tranmere_Rovers_F.C.
- 5: 298601 Bolton_Wanderers_F.C.
- 6: 333188 Barnsley_F.C.
- 7: 217730 Wolverhampton_Wanderers_F.C.
- 8: 84588 Queens_Park_Rangers_F.C.
- 9: 203434 Stoke_City_F.C.
- 10: 248083 Norwich_City_F.C.
- 11: 352843 Ipswich_Town_F.C.
- 12: 185196 Birmingham_City_F.C.
- 13: 385313 Crystal_Palace_F.C.
- 14: 451165 Oxford_United_F.C.
- 15: 236211 Huddersfield_Town_F.C.
- 16: 564301 Portsmouth_F.C.
- 17: 168007 Reading_F.C.
- 18: 165813 Manchester_City_F.C.
- 19: 33921 West_Bromwich_Albion_F.C.
- 20: 203456 Port_Vale_F.C.
- 21: 314933 Sheffield_United_F.C.
- 22: 451129 Grimsby_Town_F.C.

23: 6720 Charlton_Athletic_F.C.
24: 74291 Swindon_Town_F.C.
25: 451168 Southend_United_F.C.
26: 451131 Oldham_Athletic_A.F.C.
27: 451122 Brentford_F.C.
28: 451207 Shrewsbury_Town_F.C.
29: 451156 Bury_F.C.
30: 376725 Burnley_F.C.
31: 451121 A.F.C._Bournemouth
32: 245357 Blackpool_F.C.
33: 451126 Chesterfield_F.C.
34: 435024 Millwall_F.C.
35: 69438 Crewe_Alexandra_F.C.
36: 165278 Gillingham_F.C.
37: 429092 Preston_North_End_F.C.
38: 448599 Notts_County_F.C.
39: 241161 Bristol_Rovers_F.C.
40: 4392 Bristol_City_F.C.
41: 85099 Watford_F.C.
42: 413286 Wrexham_F.C.
43: 451136 Wycombe_Wanderers_F.C.
44: 386362 Rotherham_United_F.C.
45: 451132 Peterborough_United_F.C.
46: 359501 Walsall_F.C.
47: 451135 Stockport_County_F.C.
48: 18367 Luton_Town_F.C.
49: 451130 Hartlepool_United_F.C.
50: 367030 Wigan_Athletic_F.C.
51: 451358 Hull_City_A.F.C.
52: 451157 Carlisle_United_F.C.
53: 11228 Fulham_F.C.
54: 451167 Scunthorpe_United_F.C.
55: 946449 Scarborough_F.C.
56: 451159 Darlington_F.C.
57: 425750 Northampton_Town_F.C.
58: 451173 Barnet_F.C.
59: 451195 Chester_City_F.C.
60: 193049 Torquay_United_F.C.
61: 5624949 Cardiff_City_F.C.
62: 451169 Swansea_City_A.F.C.
63: 69455 Brighton_&_Hove_Albiaon_F.C.
64: 451201 Hereford_United_F.C.
65: 451127 Colchester_United_F.C.
66: 451166 Rochdale_A.F.C.
67: 451164 Mansfield_Town_F.C.
68: 451160 Doncaster_Rovers_F.C.
69: 451162 Leyton_Orient_F.C.
1: 431144 Bradford_City_A.F.C.
[winner: 1095011 Bradford_Bulls, pem_pos=9,
final_scores=-0.3467431, final_local_scores=-0.40980768,
pem=-3.818427, similarity=12.625435]
RECALL MISS: 431144 Bradford_City_A.F.C.]

2: 23857 Plymouth_Argyle_F.C.
[winner: 23508 Plymouth, pem_pos=0,
final_scores=-0.6110805, final_local_scores=-0.6443671,
pem=-0.45987546, similarity=-12.638612]
RECALL MISS: 23857 Plymouth_Argyle_F.C.]

3: 451170 York_City_F.C.
[winner: 992980 York_City_Knights, pem_pos=18,
final_scores=-0.64591926, final_local_scores=-0.7200541,
pem=-4.2129073, similarity=5.5123086]
RECALL MISS: 451170 York_City_F.C.]

4: 347780 Exeter_City_F.C.
[winner: 30871776 Exeter_Chiefs, pem_pos=19,
final_scores=-0.42598784, final_local_scores=-0.46853936,
pem=-3.6095665, similarity=9.736947]
RECALL MISS: 347780 Exeter_City_F.C.]

5: 243970 Cambridge_United_F.C.
[winner: 25978572 University_of_Cambridge, pem_pos=1,
final_scores=-0.89271694, final_local_scores=-0.9198106,
pem=-1.7952901, similarity=-13.596202]
RECALL MISS: 243970 Cambridge_United_F.C.]

6: 451163 Lincoln_City_F.C.
[winner: 62705 Lincoln_Motor_Company, pem_pos=2,
final_scores=-0.57769156, final_local_scores=-0.52195966,
pem=-2.4392905, similarity=1.7786381]
RECALL MISS: 451163 Lincoln_City_F.C.]

1: 9316 England
[pem_pos=1, final_scores=0.49137324,
final_local_scores=0.5563292, pem=-1.3527626,
similarity=28.048218]

2: 17867 London
[pem_pos=0, final_scores=0.6221022,
final_local_scores=0.7140864, pem=-0.09086737,
similarity=25.893625]

3: 9316 England
[pem_pos=1, final_scores=0.567149,
final_local_scores=0.63910264, pem=-1.3527626,
similarity=30.518343]

4: 69219 Tranmere_Rovers_F.C.
[pem_pos=0, final_scores=0.44919914,
final_local_scores=0.3815381, pem=-0.6002322,
similarity=18.739727]

5: 298601 Bolton_Wanderers_F.C.
[pem_pos=2, final_scores=0.5629212,
final_local_scores=0.49993223, pem=-2.2845325,
similarity=31.432333]

6: 333188 Barnsley_F.C.
[pem_pos=0, final_scores=0.39781934,
final_local_scores=0.31851003, pem=-0.7126868,
similarity=17.470388]

7: 217730 Wolverhampton_Wanderers_F.C.
[pem_pos=1, final_scores=0.81784993,
final_local_scores=0.77589285, pem=-2.274954,

```
similarity=39.615456]
8: 84588 Queens_Park_Rangers_F.C.
    [pem_pos=0, final_scores=0.4901858,
final_local_scores=0.42443955, pem=0.0, similarity=16.755836]
9: 203434 Stoke_City_F.C.
    [pem_pos=0, final_scores=0.5333428,
final_local_scores=0.4724545, pem=-0.3417574,
similarity=20.047226]
10: 248083 Norwich_City_F.C.
    [pem_pos=2, final_scores=0.8507107,
final_local_scores=0.812949, pem=-2.7278075,
similarity=43.183975]
11: 352843 Ipswich_Town_F.C.
    [pem_pos=4, final_scores=0.8306084,
final_local_scores=0.7958118, pem=-2.6931517,
similarity=42.4841]
12: 185196 Birmingham_City_F.C.
    [pem_pos=2, final_scores=0.77479583,
final_local_scores=0.732927, pem=-3.102209,
similarity=42.832012]
13: 385313 Crystal_Palace_F.C.
    [pem_pos=0, final_scores=0.5284566,
final_local_scores=0.46296197, pem=-0.51090026,
similarity=20.683777]
14: 451165 Oxford_United_F.C.
    [pem_pos=18, final_scores=0.97942454,
final_local_scores=0.9605425, pem=-4.3872933,
similarity=56.613]
15: 236211 Huddersfield_Town_F.C.
    [pem_pos=3, final_scores=0.7277896,
final_local_scores=0.6802532, pem=-2.2859883,
similarity=36.821384]
16: 564301 Portsmouth_F.C.
    [pem_pos=1, final_scores=0.74891776,
final_local_scores=0.7038595, pem=-1.2302613,
similarity=31.784637]
17: 168007 Reading_F.C.
    [pem_pos=0, final_scores=1.0149326,
final_local_scores=0.9938839, pem=-1.2007197,
similarity=40.278893]
18: 165813 Manchester_City_F.C.
    [pem_pos=0, final_scores=0.3432396,
final_local_scores=0.27308077, pem=0.0, similarity=12.238988]
19: 33921 West_Bromwich_Albion_F.C.
    [pem_pos=2, final_scores=0.6084321,
final_local_scores=0.5526286, pem=-1.5002033,
similarity=28.73959]
20: 203456 Port_Vale_F.C.
    [pem_pos=0, final_scores=0.5896825,
final_local_scores=0.5414458, pem=0.0, similarity=20.247536]
21: 314933 Sheffield_United_F.C.
    [pem_pos=0, final_scores=0.3473528,
```

```
final_local_scores=0.26339704, pem=0.0, similarity=11.950006]
22: 451129 Grimsby_Town_F.C.
    [pem_pos=2, final_scores=0.8273662,
final_local_scores=0.78916234, pem=-2.6069987,
similarity=41.817154]
23: 6720 Charlton_Athletic_F.C.
    [pem_pos=1, final_scores=0.7997357,
final_local_scores=0.76256436, pem=-2.103781,
similarity=38.286842]
24: 74291 Swindon_Town_F.C.
    [pem_pos=4, final_scores=0.8245918,
final_local_scores=0.7882531, pem=-2.489671,
similarity=41.151978]
25: 451168 Southend_United_F.C.
    [pem_pos=2, final_scores=0.740655,
final_local_scores=0.694478, pem=-1.771178,
similarity=34.446262]
26: 451131 Oldham_Athletic_A.F.C.
    [pem_pos=3, final_scores=0.7225565,
final_local_scores=0.67526215, pem=-2.8632221,
similarity=39.81153]
27: 451122 Brentford_F.C.
    [pem_pos=0, final_scores=0.7257226,
final_local_scores=0.67950934, pem=-0.36028567,
similarity=26.326918]
28: 451207 Shrewsbury_Town_F.C.
    [pem_pos=12, final_scores=0.89145166,
final_local_scores=0.86219496, pem=-3.2502756,
similarity=47.494835]
29: 451156 Bury_F.C.
    [pem_pos=0, final_scores=0.84680647,
final_local_scores=0.80894005, pem=-0.72067994,
similarity=32.149265]
30: 376725 Burnley_F.C.
    [pem_pos=1, final_scores=0.3121711,
final_local_scores=0.22803462, pem=-0.9914584,
similarity=16.286427]
31: 451121 A.F.C._Bournemouth
    [pem_pos=1, final_scores=0.58866495,
final_local_scores=0.5281282, pem=-1.0576377,
similarity=25.60171]
32: 245357 Blackpool_F.C.
    [pem_pos=1, final_scores=0.82944995,
final_local_scores=0.79039437, pem=-0.8537509,
similarity=32.31949]
33: 451126 Chesterfield_F.C.
    [pem_pos=1, final_scores=0.7671817,
final_local_scores=0.72718227, pem=-1.2678006,
similarity=32.68478]
34: 435024 Millwall_F.C.
    [pem_pos=0, final_scores=0.5020223,
final_local_scores=0.434251, pem=0.0, similarity=17.04863]
```

35: 69438 Crewe_Alexandra_F.C.
[pem_pos=3, final_scores=0.6823458,
final_local_scores=0.6335734, pem=-2.8374243,
similarity=38.427162]
36: 165278 Gillingham_F.C.
[pem_pos=0, final_scores=0.6911203,
final_local_scores=0.64284295, pem=-0.2568813,
similarity=24.67039]
37: 429092 Preston_North_End_F.C.
[pem_pos=4, final_scores=0.7758549,
final_local_scores=0.73080903, pem=-3.194944,
similarity=43.273113]
38: 448599 Notts_County_F.C.
[pem_pos=0, final_scores=0.44144374,
final_local_scores=0.3685791, pem=0.0, similarity=15.088848]
39: 241161 Bristol_Rovers_F.C.
[pem_pos=0, final_scores=0.3902154,
final_local_scores=0.3137826, pem=0.0, similarity=13.453611]
40: 4392 Bristol_City_F.C.
[pem_pos=0, final_scores=0.43450117,
final_local_scores=0.3626094, pem=0.0, similarity=14.9107]
41: 85099 Watford_F.C.
[pem_pos=1, final_scores=0.8204228,
final_local_scores=0.7823644, pem=-0.95984375,
similarity=32.656807]
42: 413286 Wrexham_F.C.
[pem_pos=1, final_scores=0.5100629,
final_local_scores=0.44961977, pem=-0.88259363,
similarity=22.306944]
43: 451136 Wycombe_Wanderers_F.C.
[pem_pos=1, final_scores=0.72062856,
final_local_scores=0.6831822, pem=-1.147249,
similarity=30.716152]
44: 386362 Rotherham_United_F.C.
[pem_pos=3, final_scores=0.49855405,
final_local_scores=0.4317515, pem=-2.6427798,
similarity=31.345884]
45: 451132 Peterborough_United_F.C.
[pem_pos=6, final_scores=0.887577,
final_local_scores=0.85755455, pem=-3.1818013,
similarity=46.983982]
46: 359501 Walsall_F.C.
[pem_pos=1, final_scores=0.6066354,
final_local_scores=0.5482924, pem=-0.61973774,
similarity=23.822083]
47: 451135 Stockport_County_F.C.
[pem_pos=4, final_scores=0.5939303,
final_local_scores=0.5397442, pem=-2.5287716,
similarity=33.948612]
48: 18367 Luton_Town_F.C.
[pem_pos=2, final_scores=0.6017751,
final_local_scores=0.5417736, pem=-2.5537767,

```
similarity=34.145153]
49: 451130 Hartlepool_United_F.C.
    [pem_pos=1, final_scores=0.64246184,
final_local_scores=0.5890216, pem=-1.75166,
similarity=31.193092]
50: 367030 Wigan_Athletic_F.C.
    [pem_pos=2, final_scores=0.525618,
final_local_scores=0.4658776, pem=-2.32737,
similarity=30.649029]
51: 451358 Hull_City_A.F.C.
    [pem_pos=11, final_scores=0.78655034,
final_local_scores=0.7448302, pem=-3.7682815,
similarity=46.809433]
52: 451157 Carlisle_United_F.C.
    [pem_pos=12, final_scores=0.90536267,
final_local_scores=0.87652797, pem=-3.5999887,
similarity=49.824356]
53: 11228 Fulham_F.C.
    [pem_pos=0, final_scores=0.5772036,
final_local_scores=0.52447885, pem=-0.3209339,
similarity=21.486496]
54: 451167 Scunthorpe_United_F.C.
    [pem_pos=1, final_scores=0.6302665,
final_local_scores=0.5753661, pem=-1.507785,
similarity=29.459352]
55: 946449 Scarborough_F.C.
    [pem_pos=2, final_scores=0.90173715,
final_local_scores=0.88890463, pem=-1.8160111,
similarity=40.49215]
56: 451159 Darlington_F.C.
    [pem_pos=1, final_scores=0.86760896,
final_local_scores=0.8361226, pem=-1.8695662,
similarity=39.208267]
57: 425750 Northampton_Town_F.C.
    [pem_pos=10, final_scores=0.9748059,
final_local_scores=0.95160705, pem=-3.2381344,
similarity=50.097042]
58: 451173 Barnet_F.C.
    [pem_pos=0, final_scores=0.7607184,
final_local_scores=0.7316562, pem=-0.6797195,
similarity=29.620213]
59: 451195 Chester_City_F.C.
    [pem_pos=1, final_scores=0.9380526,
final_local_scores=0.9248367, pem=-3.2160804,
similarity=49.17823]
60: 193049 Torquay_United_F.C.
    [pem_pos=2, final_scores=0.65759224,
final_local_scores=0.60466725, pem=-2.019306,
similarity=33.11549]
61: 5624949 Cardiff_City_F.C.
    [pem_pos=3, final_scores=0.8162766,
final_local_scores=0.7774157, pem=-3.4864018,
```

```

similarity=46.248943]
62: 451169 Swansea_City_A.F.C.
    [pem_pos=3, final_scores=0.83901495,
final_local_scores=0.8076958, pem=-2.5988693,
similarity=42.326023]
63: 69455 Brighton_&_Hove_Albion_F.C.
    [pem_pos=3, final_scores=0.92494637,
final_local_scores=0.8979421, pem=-3.1271749,
similarity=47.89216]
64: 451201 Hereford_United_F.C.
    [pem_pos=6, final_scores=1.0728877,
final_local_scores=1.0661542, pem=-2.7578042,
similarity=50.90325]
65: 451127 Colchester_United_F.C.
    [pem_pos=7, final_scores=0.87821203,
final_local_scores=0.8603695, pem=-2.8568413,
similarity=45.300804]
66: 451166 Rochdale_A.F.C.
    [pem_pos=1, final_scores=0.847102,
final_local_scores=0.81314427, pem=-0.95097107,
similarity=33.52709]
67: 451164 Mansfield_Town_F.C.
    [pem_pos=16, final_scores=0.8327957,
final_local_scores=0.7986888, pem=-4.267106,
similarity=51.129368]
68: 451160 Doncaster_Rovers_F.C.
    [pem_pos=6, final_scores=0.7312948,
final_local_scores=0.6862654, pem=-2.950685,
similarity=40.615524]
69: 451162 Leyton_Orient_F.C.
    [pem_pos=0, final_scores=0.4242428,
final_local_scores=0.35054448, pem=0.0, similarity=14.550659]

```

Sample: b'524_SOCCER'

SOCCER - [1 1 ENGLISH 1] LEAGUE RESULTS . [2 2 LONDON 2]

1996-08-27 Results of [3 3 English 3] league soccer matches

on Tuesday : Division one [4 4 Crystal Palace 4] 0 [5 5 West Bromwich 5] 0 [6 6 Ipswich 6] 1 [7 7 Grimsby 7] 1 [8 8 Oxford 8] 0 [9 9 Norwich 9] 1 [10 10 Portsmouth 10] 1 [11 11 Southend 11] 0 [12 12 Tranmere 12] 2 [13 13 Port Vale 13] 0

Postponed : [14 14 Charlton 14] v [15 15 Birmingham 15] , [16 16 Sheffield United 16] v [17 17 Huddersfield 17] Division two [18 18 Brentford 18] 2 [19 19 Gillingham 19] 0 [20 20 Bristol City 20] 5 [21 21 Luton 21] 0 [22 22 Burnley 22] 1 [23 23 Shrewsbury 23] 3 [24 24 Chesterfield 24] 1 [25 25 Walsall 25] 0 [26 26 Preston 26] 2 [27 27 Crewe 27] 1 [28 28 Rotherham 28] 1 [29 29 Blackpool 29] 2 [30 30 Stockport 30] 0 [31 31 Bourne mouth 31] 1 [32 32 Watford 32] 0 1 Plymouth 2 [33 33 Wycombe 33] 0 [34 34 Bury 34] 1 2 York 3 [35 35 Millwall 35] 2

Postponed : [36 36 Peterborough 36] v [37 37 Notts County 37] , [38 38 Wrexham 38] v [39 39 Bristol Rovers 39] Division three [40 40 Barnet 40] 3 [41 41 Brighton 41] 0 [42 42 Cardiff

42] 0 [43 43 Wigan 43] 2 [44 44 Carlisle 44] 1 [45 45 Leyton 0
rient 45] 0 [46 46 Chester 46] 2 [47 47 Swansea 47] 0 [48 48 D
arlington 48] 1 [49 49 Colchester 49] 1 3 Exeter 1 [50 50 Don
caster 50] 1 [51 51 Hartlepool 51] 2 [52 52 Mansfield 52] 2 [5
3 53 Hereford 53] 0 [54 54 Hull 54] 1 4 Lincoln 1 5 Cambridge
1 [55 55 Northampton 55] 1 [56 56 Torquay 56] 1 [57 57 Rochda
le 57] 1 [58 58 Fulham 58] 2 [59 59 Scunthorpe 59] 0 [60 60 Sc
arborough 60] 2
1: 9316 England
2: 17867 London
3: 9316 England
4: 385313 Crystal_Palace_F.C.
5: 33921 West_Bromwich_Albion_F.C.
6: 352843 Ipswich_Town_F.C.
7: 451129 Grimsby_Town_F.C.
8: 451165 Oxford_United_F.C.
9: 248083 Norwich_City_F.C.
10: 564301 Portsmouth_F.C.
11: 451168 Southend_United_F.C.
12: 69219 Tranmere_Rovers_F.C.
13: 203456 Port_Vale_F.C.
14: 6720 Charlton_Athletic_F.C.
15: 185196 Birmingham_City_F.C.
16: 314933 Sheffield_United_F.C.
17: 236211 Huddersfield_Town_F.C.
18: 451122 Brentford_F.C.
19: 165278 Gillingham_F.C.
20: 4392 Bristol_City_F.C.
21: 18367 Luton_Town_F.C.
22: 376725 Burnley_F.C.
23: 451207 Shrewsbury_Town_F.C.
24: 451126 Chesterfield_F.C.
25: 359501 Walsall_F.C.
26: 429092 Preston_North_End_F.C.
27: 69438 Crewe_Alexandra_F.C.
28: 386362 Rotherham_United_F.C.
29: 245357 Blackpool_F.C.
30: 451135 Stockport_County_F.C.
31: 451121 A.F.C._Bournemouth
32: 85099 Watford_F.C.
33: 451136 Wycombe_Wanderers_F.C.
34: 451156 Bury_F.C.
35: 435024 Millwall_F.C.
36: 451132 Peterborough_United_F.C.
37: 448599 Notts_County_F.C.
38: 413286 Wrexham_F.C.
39: 241161 Bristol_Rovers_F.C.
40: 451173 Barnet_F.C.
41: 69455 Brighton_&_Hove_Albion_F.C.
42: 5624949 Cardiff_City_F.C.
43: 367030 Wigan_Athletic_F.C.
44: 451157 Carlisle_United_F.C.

45: 451162 Leyton_Orient_F.C.
46: 451195 Chester_City_F.C.
47: 451169 Swansea_City_A.F.C.
48: 451159 Darlington_F.C.
49: 451127 Colchester_United_F.C.
50: 451160 Doncaster_Rovers_F.C.
51: 451130 Hartlepool_United_F.C.
52: 451164 Mansfield_Town_F.C.
53: 451201 Hereford_United_F.C.
54: 451358 Hull_City_A.F.C.
55: 425750 Northampton_Town_F.C.
56: 193049 Torquay_United_F.C.
57: 451166 Rochdale_A.F.C.
58: 11228 Fulham_F.C.
59: 451167 Scunthorpe_United_F.C.
60: 946449 Scarborough_F.C.
1: 23857 Plymouth_Argyle_F.C.
 [winner: 23508 Plymouth, pem_pos=0,
final_scores=-0.6201963, final_local_scores=-0.65562683,
pem=-0.45987546, similarity=-12.974626]
 RECALL MISS: 23857 Plymouth_Argyle_F.C.]
2: 451170 York_City_F.C.
 [winner: 992980 York_City_Knights, pem_pos=18,
final_scores=-0.70651996, final_local_scores=-0.78632,
pem=-4.2129073, similarity=3.5348017]
 RECALL MISS: 451170 York_City_F.C.]
3: 347780 Exeter_City_F.C.
 [winner: 30871776 Exeter_Chiefs, pem_pos=19,
final_scores=-0.40695065, final_local_scores=-0.44788176,
pem=-3.6095665, similarity=10.353412]
 RECALL MISS: 347780 Exeter_City_F.C.]
4: 451163 Lincoln_City_F.C.
 [winner: 62705 Lincoln_Motor_Company, pem_pos=2,
final_scores=-0.602679, final_local_scores=-0.54939896,
pem=-2.4392905, similarity=0.9597936]
 RECALL MISS: 451163 Lincoln_City_F.C.]
5: 243970 Cambridge_United_F.C.
 [winner: 25978572 University_of_Cambridge, pem_pos=1,
final_scores=-0.8684527, final_local_scores=-0.89406365,
pem=-1.7952901, similarity=-12.827864]
 RECALL MISS: 243970 Cambridge_United_F.C.]
1: 9316 England
 [pem_pos=1, final_scores=0.51091236,
final_local_scores=0.57691604, pem=-1.3527626,
similarity=28.66257]
2: 17867 London
 [pem_pos=0, final_scores=0.62393034,
final_local_scores=0.71612346, pem=-0.09086737,
similarity=25.954416]
3: 9316 England
 [pem_pos=1, final_scores=0.6501547,
final_local_scores=0.7290172, pem=-1.3527626,

```
similarity=33.201572]
4: 385313 Crystal_Palace_F.C.
    [pem_pos=0, final_scores=0.5064003,
final_local_scores=0.43950635, pem=-0.51090026,
similarity=19.983814]
5: 33921 West_Bromwich_Albion_F.C.
    [pem_pos=2, final_scores=0.7036516,
final_local_scores=0.6576032, pem=-1.5002033,
similarity=31.87224]
6: 352843 Ipswich_Town_F.C.
    [pem_pos=4, final_scores=0.94120175,
final_local_scores=0.91695315, pem=-2.6931517,
similarity=46.0992]
7: 451129 Grimsby_Town_F.C.
    [pem_pos=2, final_scores=0.8729983,
final_local_scores=0.83890194, pem=-2.6069987,
similarity=43.301483]
8: 451165 Oxford_United_F.C.
    [pem_pos=18, final_scores=1.0595016,
final_local_scores=1.0480413, pem=-4.3872933,
similarity=59.22414]
9: 248083 Norwich_City_F.C.
    [pem_pos=2, final_scores=0.9625613,
final_local_scores=0.93555945, pem=-2.7278075,
similarity=46.842915]
10: 564301 Portsmouth_F.C.
    [pem_pos=1, final_scores=0.74571365,
final_local_scores=0.7009545, pem=-1.2302613,
similarity=31.697945]
11: 451168 Southend_United_F.C.
    [pem_pos=2, final_scores=0.7409364,
final_local_scores=0.69455296, pem=-1.771178,
similarity=34.448498]
12: 69219 Tranmere_Rovers_F.C.
    [pem_pos=0, final_scores=0.49285477,
final_local_scores=0.42903894, pem=-0.6002322,
similarity=20.157248]
13: 203456 Port_Vale_F.C.
    [pem_pos=0, final_scores=0.5794824,
final_local_scores=0.53027785, pem=0.0, similarity=19.91426]
14: 6720 Charlton_Athletic_F.C.
    [pem_pos=1, final_scores=0.6642797,
final_local_scores=0.6151251, pem=-2.103781,
similarity=33.88696]
15: 185196 Birmingham_City_F.C.
    [pem_pos=2, final_scores=0.7611427,
final_local_scores=0.7186535, pem=-3.102209,
similarity=42.406063]
16: 314933 Sheffield_United_F.C.
    [pem_pos=0, final_scores=0.32445377,
final_local_scores=0.23867515, pem=0.0, similarity=11.2122555]
17: 236211 Huddersfield_Town_F.C.
```

[pem_pos=3, final_scores=0.684368,
final_local_scores=0.63326627, pem=-2.2859883,
similarity=35.4192]
18: 451122 Brentford_F.C.
[pem_pos=0, final_scores=0.76586825,
final_local_scores=0.7234075, pem=-0.36028567,
similarity=27.636927]
19: 165278 Gillingham_F.C.
[pem_pos=0, final_scores=0.6708784,
final_local_scores=0.6206387, pem=-0.2568813,
similarity=24.00777]
20: 4392 Bristol_City_F.C.
[pem_pos=0, final_scores=0.43647557,
final_local_scores=0.3650126, pem=0.0, similarity=14.982417]
21: 18367 Luton_Town_F.C.
[pem_pos=2, final_scores=0.70897174,
final_local_scores=0.65917176, pem=-2.5537767,
similarity=37.64855]
22: 376725 Burnley_F.C.
[pem_pos=1, final_scores=0.352987,
final_local_scores=0.27308863, pem=-0.9914584,
similarity=17.630928]
23: 451207 Shrewsbury_Town_F.C.
[pem_pos=12, final_scores=0.98988277,
final_local_scores=0.9694603, pem=-3.2502756,
similarity=50.695847]
24: 451126 Chesterfield_F.C.
[pem_pos=1, final_scores=0.89928097,
final_local_scores=0.8709944, pem=-1.2678006,
similarity=36.976418]
25: 359501 Walsall_F.C.
[pem_pos=1, final_scores=0.605578,
final_local_scores=0.5471153, pem=-0.61973774,
similarity=23.786957]
26: 429092 Preston_North_End_F.C.
[pem_pos=4, final_scores=0.8550406,
final_local_scores=0.81769174, pem=-3.194944,
similarity=45.865868]
27: 69438 Crewe_Alexandra_F.C.
[pem_pos=3, final_scores=0.8243286,
final_local_scores=0.7886268, pem=-2.8374243,
similarity=43.054264]
28: 386362 Rotherham_United_F.C.
[pem_pos=3, final_scores=0.55233634,
final_local_scores=0.49030972, pem=-2.6427798,
similarity=33.09338]
29: 245357 Blackpool_F.C.
[pem_pos=1, final_scores=0.9084013,
final_local_scores=0.8769966, pem=-0.8537509,
similarity=34.90387]
30: 451135 Stockport_County_F.C.
[pem_pos=4, final_scores=0.631773,

```
final_local_scores=0.5807734, pem=-2.5287716,
similarity=35.173004]
31: 451121 A.F.C._Bournemouth
    [pem_pos=1, final_scores=0.8098231,
final_local_scores=0.769795, pem=-1.0576377,
similarity=32.81353]
32: 85099 Watford_F.C.
    [pem_pos=1, final_scores=0.8061324,
final_local_scores=0.76715666, pem=-0.95984375,
similarity=32.202976]
33: 451136 Wycombe_Wanderers_F.C.
    [pem_pos=1, final_scores=0.8064423,
final_local_scores=0.7768175, pem=-1.147249,
similarity=33.510414]
34: 451156 Bury_F.C.
    [pem_pos=0, final_scores=0.7706565,
final_local_scores=0.725737, pem=-0.72067994,
similarity=29.66632]
35: 435024 Millwall_F.C.
    [pem_pos=0, final_scores=0.5748169,
final_local_scores=0.513926, pem=0.0, similarity=19.42629]
36: 451132 Peterborough_United_F.C.
    [pem_pos=6, final_scores=1.0357752,
final_local_scores=1.019532, pem=-3.1818013,
similarity=51.81771]
37: 448599 Notts_County_F.C.
    [pem_pos=0, final_scores=0.4655376,
final_local_scores=0.3948797, pem=0.0, similarity=15.873712]
38: 413286 Wrexham_F.C.
    [pem_pos=1, final_scores=0.61222905,
final_local_scores=0.56104016, pem=-0.88259363,
similarity=25.63195]
39: 241161 Bristol_Rovers_F.C.
    [pem_pos=0, final_scores=0.41716045,
final_local_scores=0.34303856, pem=0.0, similarity=14.326668]
40: 451173 Barnet_F.C.
    [pem_pos=0, final_scores=0.70242655,
final_local_scores=0.6679345, pem=-0.6797195,
similarity=27.71863]
41: 69455 Brighton_&Hove_Albion_F.C.
    [pem_pos=3, final_scores=1.1096951,
final_local_scores=1.0998721, pem=-3.1271749,
similarity=53.91815]
42: 5624949 Cardiff_City_F.C.
    [pem_pos=3, final_scores=0.8766075,
final_local_scores=0.84385484, pem=-3.4864018,
similarity=48.23162]
43: 367030 Wigan_Athletic_F.C.
    [pem_pos=2, final_scores=0.61898625,
final_local_scores=0.56881267, pem=-2.32737,
similarity=33.720818]
44: 451157 Carlisle_United_F.C.
```

[pem_pos=12, final_scores=0.8986047,
final_local_scores=0.8688534, pem=-3.5999887,
similarity=49.59533]
45: 451162 Leyton_Orient_F.C.
[pem_pos=0, final_scores=0.45505172,
final_local_scores=0.38417792, pem=0.0, similarity=15.554349]
46: 451195 Chester_City_F.C.
[pem_pos=1, final_scores=0.8820669,
final_local_scores=0.86326915, pem=-3.2160804,
similarity=47.34093]
47: 451169 Swansea_City_A.F.C.
[pem_pos=3, final_scores=1.0221138,
final_local_scores=1.0083277, pem=-2.5988693,
similarity=48.31328]
48: 451159 Darlington_F.C.
[pem_pos=1, final_scores=0.89120525,
final_local_scores=0.86152, pem=-1.8695662,
similarity=39.96618]
49: 451127 Colchester_United_F.C.
[pem_pos=7, final_scores=1.0711422,
final_local_scores=1.0705469, pem=-2.8568413,
similarity=51.572914]
50: 451160 Doncaster_Rovers_F.C.
[pem_pos=6, final_scores=0.7212537,
final_local_scores=0.6749938, pem=-2.950685,
similarity=40.27916]
51: 451130 Hartlepool_United_F.C.
[pem_pos=1, final_scores=0.79173225,
final_local_scores=0.7517111, pem=-1.75166,
similarity=36.048073]
52: 451164 Mansfield_Town_F.C.
[pem_pos=16, final_scores=0.9435647,
final_local_scores=0.9193478, pem=-4.267106,
similarity=54.730072]
53: 451201 Hereford_United_F.C.
[pem_pos=6, final_scores=1.2066247,
final_local_scores=1.2119801, pem=-2.7578042,
similarity=55.254982]
54: 451358 Hull_City_A.F.C.
[pem_pos=11, final_scores=0.8277088,
final_local_scores=0.790072, pem=-3.7682815,
similarity=48.15954]
55: 425750 Northampton_Town_F.C.
[pem_pos=10, final_scores=0.9817472,
final_local_scores=0.9590271, pem=-3.2381344,
similarity=50.318474]
56: 193049 Torquay_United_F.C.
[pem_pos=2, final_scores=0.70467216,
final_local_scores=0.65585035, pem=-2.019306,
similarity=34.642895]
57: 451166 Rochdale_A.F.C.
[pem_pos=1, final_scores=0.8388435,

```

final_local_scores=0.8037879, pem=-0.95097107,
similarity=33.247875]
58: 11228 Fulham_F.C.
    [pem_pos=0, final_scores=0.5459128,
final_local_scores=0.49138707, pem=-0.3209339,
similarity=20.498972]
59: 451167 Scunthorpe_United_F.C.
    [pem_pos=1, final_scores=0.7297846,
final_local_scores=0.68377215, pem=-1.507785,
similarity=32.694405]
60: 946449 Scarborough_F.C.
    [pem_pos=2, final_scores=1.0075352,
final_local_scores=1.0039537, pem=-1.8160111,
similarity=43.92544]

```

Sample: b'525_CRICKET'

CRICKET - [1 1 ENGLAND 1] AND [2 2 PAKISTAN 2] TEST
AVERAGES . [3 3 LONDON 3] 1996-08-27 [4 4 England 4] and [5 5
Pakistan 5] Test averages after their three-match series
which ended on Monday : [6 6 England 6] Batting (tabulated
under matches , innings , not outs , runs , highest score ,
average) : [7 7 Alec Stewart 7] 3 5 0 396 170 79.20 [8 8 John
Crawley 8] 2 3 0 178 106 59.33 [9 9 Nick Knight 9] 3 5 0 190
113 38.00 [10 10 Nasser Hussain 10] 2 3 0 111 51 37.00 [11 11
Mike Atherton 11] 3 5 0 162 64 32.40 [12 12 Graham Thorpe 12]
3 5 0 159 77 31.80 Jack Russell 2 3 1 51 41no 25.50 [13 13 Ian
Salisbury 13] 2 4 1 50 40 16.66 [14 14 Mark Ealham 14] 1 2 0
30 25 15.00 [15 15 Dominic Cork 15] 3 5 0 58 26 11.60 [16 16 R
obert Croft 16] 1 2 1 11 6 11.00 [17 17 Simon Brown 17] 1 2 1
11 10no 11.00 [18 18 Alan Mullally 18] 3 5 1 39 24 9.75 [19 19
Chris Lewis 19] 2 3 0 18 9 6.00 [20 20 Andy Caddick 20] 1 1 0
4 4 4.00 [21 21 Graeme Hick 21] 1 2 0 8 4 4.00 Bowling
(tabulated under overs , maidens , runs , wickets ,
average) : [22 22 Atherton 22] 7 1 20 1 20.00 [23 23 Caddick
23] 57.2 10 165 6 27.50 [24 24 Cork 24] 131 23 434 12 36.16 [2
5 25 Mullally 25] 150.3 36 377 10 37.70 [26 26 Hick 26] 13 2
42 1 42.00 [27 27 Croft 27] 47.4 10 125 2 62.50 Brown 33 4
138 2 69.00 [28 28 Ealham 28] 37 8 81 1 81.00 [29 29 Salisbury
29] 61.2 8 221 2 110.50 Lewis 71 10 264 1 264.00 [30 30 Thorpe
30] 13 4 19 0 - [31 31 Pakistan 31] Batting : [32 32 Moin Khan
32] 2 3 1 158 105 79.00 [33 33 Ijaz Ahmed 33] 3 6 1 344 141
68.80 [34 34 Salim Malik 34] 3 5 2 195 100no 65.00 [35 35 Inz
amam-ul-Haq 35] 3 5 0 320 148 64.00 [36 36 Saeed Anwar 36] 3
6 0 362 176 60.33 [37 37 Rashid Latif 37] 1 1 0 45 45 45.00
Aamir Sohail 2 3 1 77 46 38.50 [38 38 Asif Mujtaba 38] 2 3 0
90 51 30.00 [39 39 Wasim Akram 39] 3 5 1 98 40 24.50 [40 40 Sh
adab Kabir 40] 2 4 0 87 35 21.75 [41 41 Mushtaq Ahmed 41] 3 5
1 44 20 11.00 [42 42 Waqar Younis 42] 3 3 1 11 7 5.50 [43 43 A
ta-ur-Rehman 43] 2 2 2 10 10no - [44 44 Mohammad Akram 44] 1
0 0 0 0 - Bowling : [45 45 Mushtaq Ahmed 45] 195 52 447 17
26.29 [46 46 Waqar Younis 46] 125 25 431 16 26.93 [47 47 Wasim
Akram 47] 128 29 350 11 31.81 [48 48 Ata-ur-Rehman 48] 48.4 6

173 5 34.60 [49 49 Mohammad_Akram 49] 22 4 71 1 71.00 Aamir
Sohail 11 3 24 0 - [50 50 Asif_Mujtaba 50] 7 5 6 0 - [51 51 Sa
lim_Malik 51] 1 0 1 0 - [52 52 Shadab_Kabir 52] 1 0 9 0 -
1: 759125 England_cricket_team
2: 759295 Pakistan_national_cricket_team
3: 17867 London
4: 759125 England_cricket_team
5: 759295 Pakistan_national_cricket_team
6: 759125 England_cricket_team
7: 896414 Alec_Stewart
8: 2930949 John_Crawley
9: 994718 Nick_Knight
10: 682925 Nasser_Hussain
11: 159024 Michael_Atherton
12: 397259 Graham_Thorpe
13: 2416655 Ian_Salisbury
14: 3311473 Mark_Ealham
15: 2036840 Dominic_Cork
16: 1160680 Robert_Croft
17: 3733951 Simon_Brown_(cricketer)
18: 2222543 Alan_Mullally
19: 1732971 Chris_Lewis_(cricketer)
20: 896396 Andrew_Caddick
21: 668698 Graeme_Hick
22: 159024 Michael_Atherton
23: 896396 Andrew_Caddick
24: 2036840 Dominic_Cork
25: 2222543 Alan_Mullally
26: 668698 Graeme_Hick
27: 1160680 Robert_Croft
28: 3311473 Mark_Ealham
29: 2416655 Ian_Salisbury
30: 397259 Graham_Thorpe
31: 759295 Pakistan_national_cricket_team
32: 554132 Moin_Khan
33: 938863 Ijaz_Ahmed_(cricketer)
34: 938861 Saleem_Malik
35: 896403 Inzamam-ul-Haq
36: 894437 Saeed_Anwar
37: 894438 Rashid_Latif
38: 3976606 Asif_Mujtaba
39: 501393 Wasim_Akram
40: 3976883 Shadab_Kabir
41: 896405 Mushtaq_Ahmed
42: 894433 Waqar_Younis
43: 3976751 Ata-ur-Rehman
44: 1958830 Mohammad_Akram
45: 896405 Mushtaq_Ahmed
46: 894433 Waqar_Younis
47: 501393 Wasim_Akram
48: 3976751 Ata-ur-Rehman
49: 1958830 Mohammad_Akram

```
50: 3976606 Asif_Mujtaba
51: 938861 Saleem_Malik
52: 3976883 Shadab_Kabir
1: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.6916532,
final_local_scores=0.6681058, pem=-3.3054674,
similarity=42.002968]
2: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.73790306,
final_local_scores=0.7252193, pem=-2.7556365,
similarity=40.71728]
3: 17867 London
    [pem_pos=0, final_scores=0.4893086,
final_local_scores=0.61590475, pem=-0.09086737,
similarity=22.96369]
4: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.7087344,
final_local_scores=0.6867644, pem=-3.3054674,
similarity=42.55978]
5: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.559241,
final_local_scores=0.5300582, pem=-2.7556365,
similarity=34.893288]
6: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.67537636,
final_local_scores=0.65032583, pem=-3.3054674,
similarity=41.472378]
7: 896414 Alec_Stewart
    [pem_pos=0, final_scores=0.80364305,
final_local_scores=0.8139835, pem=0.0, similarity=28.380604]
8: 2930949 John_Crawley
    [pem_pos=0, final_scores=0.5896955,
final_local_scores=0.57620376, pem=0.0, similarity=21.284782]
9: 994718 Nick_Knight
    [pem_pos=1, final_scores=0.5265592,
final_local_scores=0.56709963, pem=-0.55008924,
similarity=24.004568]
10: 682925 Nasser_Hussain
    [pem_pos=0, final_scores=0.5011502,
final_local_scores=0.48462743, pem=0.0, similarity=18.551962]
11: 159024 Michael_Atherton
    [pem_pos=0, final_scores=0.71837205,
final_local_scores=0.7138466, pem=0.0, similarity=25.39232]
12: 397259 Graham_Thorpe
    [pem_pos=0, final_scores=0.61508733,
final_local_scores=0.6073476, pem=0.0, similarity=22.214176]
13: 2416655 Ian_Salisbury
    [pem_pos=0, final_scores=0.58216715,
final_local_scores=0.5915216, pem=0.0, similarity=21.741898]
14: 3311473 Mark_Ealham
    [pem_pos=0, final_scores=0.50510234,
final_local_scores=0.5125963, pem=0.0, similarity=19.38661]
```

```
15: 2036840 Dominic_Cork
    [pem_pos=0, final_scores=0.5473091,
final_local_scores=0.521427, pem=0.0, similarity=19.650133]
16: 1160680 Robert_Croft
    [pem_pos=0, final_scores=0.8765331,
final_local_scores=0.9064277, pem=0.0, similarity=31.139322]
17: 3733951 Simon_Brown_(cricketer)
    [pem_pos=2, final_scores=0.34030527,
final_local_scores=0.38351744, pem=-1.5092784,
similarity=23.742329]
18: 2222543 Alan_Mullally
    [pem_pos=0, final_scores=0.697532,
final_local_scores=0.71300584, pem=0.0, similarity=25.36723]
19: 1732971 Chris_Lewis_(cricketer)
    [pem_pos=1, final_scores=0.464715,
final_local_scores=0.4961837, pem=-0.7619004,
similarity=23.040157]
20: 896396 Andrew_Caddick
    [pem_pos=0, final_scores=0.5138592,
final_local_scores=0.49401742, pem=0.0, similarity=18.832178]
21: 668698 Graeme_Hick
    [pem_pos=0, final_scores=0.90127105,
final_local_scores=0.9035954, pem=0.0, similarity=31.0548]
22: 159024 Michael_Atherton
    [pem_pos=0, final_scores=1.0865394,
final_local_scores=1.1183884, pem=0.0, similarity=37.464653]
23: 896396 Andrew_Caddick
    [pem_pos=0, final_scores=0.61700743,
final_local_scores=0.6066912, pem=0.0, similarity=22.194588]
24: 2036840 Dominic_Cork
    [pem_pos=0, final_scores=1.2195643,
final_local_scores=1.2579255, pem=0.0, similarity=41.62872]
25: 2222543 Alan_Mullally
    [pem_pos=0, final_scores=0.7319556,
final_local_scores=0.7537865, pem=0.0, similarity=26.584206]
26: 668698 Graeme_Hick
    [pem_pos=0, final_scores=0.7022994,
final_local_scores=0.68823636, pem=0.0, similarity=24.62806]
27: 1160680 Robert_Croft
    [pem_pos=0, final_scores=1.0171996,
final_local_scores=1.0631881, pem=0.0, similarity=35.817364]
28: 3311473 Mark_Ealham
    [pem_pos=0, final_scores=0.37292928,
final_local_scores=0.3682173, pem=0.0, similarity=15.078052]
29: 2416655 Ian_Salisbury
    [pem_pos=0, final_scores=0.6788494,
final_local_scores=0.6971323, pem=0.0, similarity=24.893532]
30: 397259 Graham_Thorpe
    [pem_pos=0, final_scores=0.85624176,
final_local_scores=0.8733848, pem=0.0, similarity=30.153255]
31: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7184487,
```

```
final_local_scores=0.70396835, pem=-2.7556365,
similarity=40.08311]
32: 554132 Moin_Khan
    [pem_pos=0, final_scores=0.6203322,
final_local_scores=0.6576598, pem=0.0, similarity=23.715593]
33: 938863 Ijaz_Ahmed_(cricketer)
    [pem_pos=0, final_scores=0.5734996,
final_local_scores=0.5709599, pem=0.0, similarity=21.128296]
34: 938861 Saleem_Malik
    [pem_pos=0, final_scores=0.6865172,
final_local_scores=0.7059046, pem=0.0, similarity=25.155313]
35: 896403 Inzamam-ul-Haq
    [pem_pos=0, final_scores=0.87756187,
final_local_scores=0.88668114, pem=0.0, similarity=30.550045]
36: 894437 Saeed_Anwar
    [pem_pos=0, final_scores=0.7046981,
final_local_scores=0.7122547, pem=0.0, similarity=25.344814]
37: 894438 Rashid_Latif
    [pem_pos=0, final_scores=0.50343204,
final_local_scores=0.531078, pem=0.0, similarity=19.938139]
38: 3976606 Asif_Mujtaba
    [pem_pos=0, final_scores=0.5884432,
final_local_scores=0.6041253, pem=0.0, similarity=22.118017]
39: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.71652776,
final_local_scores=0.70852226, pem=0.0, similarity=25.23343]
40: 3976883 Shadab_Kabir
    [pem_pos=0, final_scores=0.41286206,
final_local_scores=0.52253693, pem=0.0, similarity=19.683258]
41: 896405 Mushtaq_Ahmed
    [pem_pos=0, final_scores=0.56452626,
final_local_scores=0.5541218, pem=0.0, similarity=20.625813]
42: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.6424305,
final_local_scores=0.63590884, pem=0.0, similarity=23.066502]
43: 3976751 Ata-ur-Rehman
    [pem_pos=0, final_scores=0.5881964,
final_local_scores=0.65611, pem=0.0, similarity=23.669344]
44: 1958830 Mohammad_Akram
    [pem_pos=0, final_scores=0.56890965,
final_local_scores=0.5831178, pem=0.0, similarity=21.49111]
45: 896405 Mushtaq_Ahmed
    [pem_pos=0, final_scores=0.5500659,
final_local_scores=0.53832597, pem=0.0, similarity=20.154434]
46: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.61220294,
final_local_scores=0.60288996, pem=0.0, similarity=22.081152]
47: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.690391,
final_local_scores=0.67997175, pem=0.0, similarity=24.381426]
48: 3976751 Ata-ur-Rehman
    [pem_pos=0, final_scores=0.6033356,
```

```
final_local_scores=0.6726473, pem=0.0, similarity=24.16285]
49: 1958830 Mohammad_Akram
    [pem_pos=0, final_scores=0.57558775,
final_local_scores=0.59041256, pem=0.0, similarity=21.708801]
50: 3976606 Asif_Mujtaba
    [pem_pos=0, final_scores=0.67464715,
final_local_scores=0.69828993, pem=0.0, similarity=24.928078]
51: 938861 Saleem_Malik
    [pem_pos=0, final_scores=0.50795084,
final_local_scores=0.5108479, pem=0.0, similarity=19.334435]
52: 3976883 Shadab_Kabir
    [pem_pos=0, final_scores=0.37801522,
final_local_scores=0.48447204, pem=0.0, similarity=18.547325]
```

Sample: b'526_CRICKET'

CRICKET - [1 1 GOOCH 1] TO PLAY ANOTHER SEASON FOR [1 1 ESSEX 1] . [2 2 LONDON 2] 1996-08-27 [3 3 Graham Gooch 3] , the 43-year-old former [4 4 England 4] captain , is to continue playing county cricket for at least another season , his club [5 5 Essex 5] announced on Tuesday . Opener [6 6 Gooch 6] 's decision comes towards the end of a season in which he has underlined his consistency by becoming the leading scorer in [7 7 Essex 7] 's history , beating [8 8 Keith Fletcher 8] 's aggregate of 29,434 . [9 9 Gooch 9] , who retired from test cricket after the 1994-95 tour of [10 10 Australia 10] but is now an [11 11 England 11] selector , is seventh in this season 's first-class averages with 1,429 runs at 64.95 , having hit five centuries and one double century . [12 12 Essex 12] secretary-general manager Peter Edwards said : " He is a remarkable batsman and still the best in this country . No-one will argue with that . You just have to look at his record to appreciate that fact . "

```
1: 413176 Graham_Gooch
2: 17867 London
3: 413176 Graham_Gooch
4: 759125 England_cricket_team
5: 1622252 Essex_County_Cricket_Club
6: 413176 Graham_Gooch
7: 1622252 Essex_County_Cricket_Club
8: 1354358 Keith_Fletcher
9: 413176 Graham_Gooch
10: 4689264 Australia
11: 759125 England_cricket_team
12: 1622252 Essex_County_Cricket_Club
1: 40100758 Essex
    [winner: 1622252 Essex_County_Cricket_Club, pem_pos=1,
final_scores=0.07302518, final_local_scores=0.004875183,
pem=-2.3459642, similarity=16.992916]
    gt: 40100758 Essex, pem_pos=0, final_scores=0.046453074,
final_local_scores=0.058043897, pem=-0.576716,
similarity=8.95813 ]
1: 413176 Graham_Gooch
```

```
[pem_pos=0, final_scores=1.1718435,
final_local_scores=1.1866808, pem=0.0, similarity=39.502636]
2: 17867 London
    [pem_pos=0, final_scores=0.6350146,
final_local_scores=0.722588, pem=-0.09086737,
similarity=26.147331]
3: 413176 Graham_Gooch
    [pem_pos=0, final_scores=1.0218359,
final_local_scores=1.0181351, pem=0.0, similarity=34.472897]
4: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.6631368,
final_local_scores=0.6271315, pem=-3.3054674,
similarity=40.780212]
5: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.76815933,
final_local_scores=0.76208377, pem=-2.3459642,
similarity=39.58953]
6: 413176 Graham_Gooch
    [pem_pos=0, final_scores=1.072396,
final_local_scores=1.0780497, pem=0.0, similarity=36.260864]
7: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.91595715,
final_local_scores=0.9235304, pem=-2.3459642,
similarity=44.407417]
8: 1354358 Keith_Fletcher
    [pem_pos=0, final_scores=0.84199196,
final_local_scores=0.9096852, pem=0.0, similarity=31.236534]
9: 413176 Graham_Gooch
    [pem_pos=0, final_scores=1.2932827,
final_local_scores=1.3193347, pem=0.0, similarity=43.461296]
10: 4689264 Australia
    [pem_pos=0, final_scores=0.54922676,
final_local_scores=0.6605893, pem=-0.17168477,
similarity=24.736664]
11: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.58660895,
final_local_scores=0.5435365, pem=-3.3054674,
similarity=38.285572]
12: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.66261894,
final_local_scores=0.6467968, pem=-2.3459642,
similarity=36.14914]
1: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.07302518,
final_local_scores=0.004875183, pem=-2.3459642,
similarity=16.992916]
```

Sample: b'527_SOCCER'

SOCCER - [1 1 ROMANIA 1] CLUB BOSS BANNED FOR HEADBUTT . [2 2
BUCHAREST 2] 1996-08-27 The Romanian Soccer Federation has
banned first division club [3 3 Jiul Petrosani 3] 's
president [4 4 Miron Cozma 4] for two years for headbutting a

visiting team player , a federation statement said . [5 5 Romania 5] 's soccer bosses also fined [6 6 Cozma 6] , a well-known miners ' union leader , 10 million lei (\$ 3000) for the half-time attack on [7 7 Dinamo Bucharest 7] 's [8 8 Danut Lupu 8] last Sunday . Miners led by [9 9 Cozma 9] rioted in [10 10 Bucharest 10] in 1990 and 1991 , bringing down the reformist government of premier [11 11 Petre Roman 11] . [12 12 Cozma 12] is awaiting trial for assault and criminal damage in a bar in his home town of Petrosan , 300 kms west of [13 13 Bucharest 13] . The attack on [14 14 Lupu 14] came during a tunnel skirmish between opposing players as they left the field . " [15 15 Cozma 15] 's blow was not too painful because I 'm a tall man , " [16 16 Lupu 16] told [17 17 Reuters 17] on Tuesday . [18 18 Lupu 18] is one of the tallest players in [19 19 Romania 19] 's first division , towering over [20 20 Cozma 20] by some 17 cms , [21 21 Jiul Petrosani 21] , promoted to the first division this year , won the league game 1-0 . [22 22 Cozma 22] is barred from taking part in any official soccer activity during the ban .

1: 25445 Romania
2: 36877 Bucharest
3: 3574206 CS_Jiul_Petroșani
4: 1285207 Miron_Cozma
5: 25445 Romania
6: 1285207 Miron_Cozma
7: 861358 FC_Dinamo_București
8: 5210442 Dănuț_Lupu
9: 1285207 Miron_Cozma
10: 36877 Bucharest
11: 502696 Petre_Roman
12: 1285207 Miron_Cozma
13: 36877 Bucharest
14: 5210442 Dănuț_Lupu
15: 1285207 Miron_Cozma
16: 5210442 Dănuț_Lupu
17: 18998750 Reuters
18: 5210442 Dănuț_Lupu
19: 25445 Romania
20: 1285207 Miron_Cozma
21: 3574206 CS_Jiul_Petroșani
22: 1285207 Miron_Cozma
1: 25445 Romania
[pem_pos=0, final_scores=0.78009456, final_local_scores=0.84164387, pem=-0.21241479, similarity=30.36119]
2: 36877 Bucharest
[pem_pos=0, final_scores=0.46254718, final_local_scores=0.47514492, pem=-0.01781679, similarity=18.365875]
3: 3574206 CS_Jiul_Petroșani
[pem_pos=0, final_scores=0.36361623, final_local_scores=0.40579432, pem=0.0, similarity=16.199425]

```
4: 1285207 Miron_Cozma
    [pem_pos=0, final_scores=0.5534846,
final_local_scores=0.53914434, pem=0.0, similarity=20.178856]
5: 25445 Romania
    [pem_pos=0, final_scores=0.5535813,
final_local_scores=0.5942125, pem=-0.21241479,
similarity=22.977343]
6: 1285207 Miron_Cozma
    [pem_pos=1, final_scores=0.92214257,
final_local_scores=0.94727486, pem=0.0, similarity=32.358284]
7: 861358 FC_Dinamo_Bucureşti
    [pem_pos=0, final_scores=0.24850242,
final_local_scores=0.27207208, pem=0.0, similarity=12.208886]
8: 5210442 Dănuț_Lupu
    [pem_pos=0, final_scores=0.5553494,
final_local_scores=0.6243411, pem=0.0, similarity=22.721296]
9: 1285207 Miron_Cozma
    [pem_pos=1, final_scores=0.87211174,
final_local_scores=0.8926237, pem=0.0, similarity=30.727385]
10: 36877 Bucharest
    [pem_pos=0, final_scores=0.46137625,
final_local_scores=0.47386587, pem=-0.01781679,
similarity=18.327705]
11: 502696 Petre_Roman
    [pem_pos=0, final_scores=0.48365426,
final_local_scores=0.5603343, pem=0.0, similarity=20.811207]
12: 1285207 Miron_Cozma
    [pem_pos=1, final_scores=0.85074097,
final_local_scores=0.86927944, pem=0.0, similarity=30.030745]
13: 36877 Bucharest
    [pem_pos=0, final_scores=0.33152676,
final_local_scores=0.33202502, pem=-0.01781679,
similarity=14.0948925]
14: 5210442 Dănuț_Lupu
    [pem_pos=0, final_scores=0.7228951,
final_local_scores=0.8183055, pem=0.0, similarity=28.50958]
15: 1285207 Miron_Cozma
    [pem_pos=1, final_scores=0.8268799,
final_local_scores=0.8432148, pem=0.0, similarity=29.252924]
16: 5210442 Dănuț_Lupu
    [pem_pos=0, final_scores=0.72557193,
final_local_scores=0.8212296, pem=0.0, similarity=28.596842]
17: 18998750 Reuters
    [pem_pos=0, final_scores=0.25181872,
final_local_scores=0.39142692, pem=0.0, similarity=15.770674]
18: 5210442 Dănuț_Lupu
    [pem_pos=0, final_scores=0.7435147,
final_local_scores=0.8408294, pem=0.0, similarity=29.181738]
19: 25445 Romania
    [pem_pos=0, final_scores=0.58792657,
final_local_scores=0.6317295, pem=-0.21241479,
similarity=24.096926]
```

```
20: 1285207 Miron_Cozma
    [pem_pos=1, final_scores=0.6525139,
final_local_scores=0.65274644, pem=0.0, similarity=23.568968]
21: 3574206 CS_Jiul_Petroșani
    [pem_pos=0, final_scores=0.35824472,
final_local_scores=0.39992678, pem=0.0, similarity=16.024326]
22: 1285207 Miron_Cozma
    [pem_pos=1, final_scores=0.79654926,
final_local_scores=0.81008327, pem=0.0, similarity=28.264214]
```

Sample: b'528_SQUASH'

SQUASH - HONG KONG OPEN FIRST ROUND RESULTS . [1 1 HONG KONG
1] 1996-08-27 First round results in the Hong Kong Open
squash tournament on Tuesday (prefix denotes seeding) : 1 -
[2 2 Jansher Khan 2] (Pakistnn) beat Jackie Lee ([3 3 Hong
Kong 3]) 15-8 15-8 15-6 3 - [4 4 Brett Martin 4] ([5 5 Aust
ralia 5]) beat [6 6 David Evans 6] ([7 7 Wales 7]) 14-17
15-1 13-15 17-14 15-12 [8 8 Mark Cairns 8] ([9 9 England
9]) beat 6 - [10 10 Del Harris 10] ([11 11 England 11])
15-12 7-15 15-6 15-12 Anthony Hill ([12 12 Australia 12])
beat 8 - [13 13 Mark Chaloner 13] ([14 14 England 14])
15-11 17-16 17-16 Simon Frenz ([15 15 Germany 15]) beat [16
16 Martin Heath 16] ([17 17 Scotland 17]) 12-15 15-6 15-4
12-15 15-14 [18 18 Joseph Kneipp 18] ([19 19 Australia 19])
beat Ahmed Faizy ([20 20 Egypt 20]) 15-8 12-15 15-14 15-9
Mir Zaman Gul ([21 21 Pakistan 21]) beat Stephen Meads ([22
22 England 22]) 10-15 15-12 15-10 15-3 Dan Jensen ([23 23 Au
stralia 23]) beat Anders Thoren ([24 24 Sweden 24]) 8-15
15-12 10-15 15-5 15-11

```
1: 13404 Hong_Kong
2: 2525758 Jansher_Khan
3: 13404 Hong_Kong
4: 3428114 Brett_Martin
5: 4689264 Australia
6: 3671790 David_Evans_(squash_player)
7: 69894 Wales
8: 7555328 Mark_Cairns_(squash_player)
9: 9316 England
10: 3428265 Del_Harris_(squash_player)
11: 9316 England
12: 4689264 Australia
13: 7434417 Mark_Chaloner
14: 9316 England
15: 11867 Germany
16: 8230495 Martin_Heath
17: 26994 Scotland
18: 6849383 Joseph_Kneipp
19: 4689264 Australia
20: 8087628 Egypt
21: 23235 Pakistan
22: 9316 England
23: 4689264 Australia
```

```
24: 5058739 Sweden
1: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.58072484,
final_local_scores=0.58195746, pem=-0.22402541,
similarity=22.67477]
2: 2525758 Jansher_Khan
    [pem_pos=0, final_scores=0.54349047,
final_local_scores=0.67714506, pem=0.0, similarity=24.297071]
3: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.53971016,
final_local_scores=0.53715515, pem=-0.22402541,
similarity=21.33778]
4: 3428114 Brett_Martin
    [pem_pos=0, final_scores=0.7005138,
final_local_scores=0.84152997, pem=0.0, similarity=29.202646]
5: 4689264 Australia
    [pem_pos=0, final_scores=0.75083226,
final_local_scores=0.7472639, pem=-0.17168477,
similarity=27.323208]
6: 3671790 David_Evans_(squash_player)
    [pem_pos=5, final_scores=0.32808715,
final_local_scores=0.42515266, pem=-2.1692557,
similarity=28.573866]
7: 69894 Wales
    [pem_pos=0, final_scores=0.5025386,
final_local_scores=0.49550533, pem=-0.37222588,
similarity=20.900803]
8: 7555328 Mark_Cairns_(squash_player)
    [pem_pos=0, final_scores=0.46751612,
final_local_scores=0.5818601, pem=-0.18232158,
similarity=22.445072]
9: 9316 England
    [pem_pos=0, final_scores=0.5180494,
final_local_scores=0.48476714, pem=-0.19053856,
similarity=19.59231]
10: 3428265 Del_Harris_(squash_player)
    [pem_pos=1, final_scores=0.48179877,
final_local_scores=0.6076656, pem=-0.62365466,
similarity=25.615198]
11: 9316 England
    [pem_pos=0, final_scores=0.54240423,
final_local_scores=0.51137114, pem=-0.19053856,
similarity=20.386227]
12: 4689264 Australia
    [pem_pos=0, final_scores=0.70195395,
final_local_scores=0.6938718, pem=-0.17168477,
similarity=25.729881]
13: 7434417 Mark_Chaloner
    [pem_pos=0, final_scores=0.648783,
final_local_scores=0.7857523, pem=0.0, similarity=27.538128]
14: 9316 England
    [pem_pos=0, final_scores=0.5472521,
```

```
final_local_scores=0.51666665, pem=-0.19053856,
similarity=20.544254]
15: 11867 Germany
    [pem_pos=0, final_scores=0.87346524,
final_local_scores=0.9158241, pem=-0.15529114,
similarity=32.26423]
16: 8230495 Martin_Heath
    [pem_pos=0, final_scores=0.6163232,
final_local_scores=0.7325502, pem=0.0, similarity=25.950472]
17: 26994 Scotland
    [pem_pos=0, final_scores=0.46990603,
final_local_scores=0.4608925, pem=-0.22905664,
similarity=19.08931]
18: 6849383 Joseph_Kneipp
    [pem_pos=0, final_scores=0.693015,
final_local_scores=0.8222796, pem=0.0, similarity=28.628176]
19: 4689264 Australia
    [pem_pos=0, final_scores=0.77583736,
final_local_scores=0.77457815, pem=-0.17168477,
similarity=28.138319]
20: 8087628 Egypt
    [pem_pos=0, final_scores=0.5762344,
final_local_scores=0.6090644, pem=-0.27480945,
similarity=23.759867]
21: 23235 Pakistan
    [pem_pos=0, final_scores=0.64365226,
final_local_scores=0.6811789, pem=-0.21944217,
similarity=25.610811]
22: 9316 England
    [pem_pos=0, final_scores=0.5471623,
final_local_scores=0.51656854, pem=-0.19053856,
similarity=20.541327]
23: 4689264 Australia
    [pem_pos=0, final_scores=0.7082917,
final_local_scores=0.70079476, pem=-0.17168477,
similarity=25.936476]
24: 5058739 Sweden
    [pem_pos=0, final_scores=0.77783376,
final_local_scores=0.81504166, pem=-0.2135508,
similarity=29.573503]
```

Sample: b'529_TENNIS'

TENNIS - [1 1 EDBERG 1] EXTENDS GRAND SLAM RUN , TOPPLES [2 2 WIMBLEDON 2] CHAMP . Larry Fine [3 3 NEW YORK 3] 1996-08-27 [4 4 Stefan Edberg 4] produced some of his vintage best on Tuesday to extend his grand run at the [5 5 Grand Slams 5] by toppling [6 6 Wimbledon 6] champion [7 7 Richard Krajicek 7] in straight sets at the [8 8 U.S. Open 8] . [9 9 Edberg 9] , competing in the 54th consecutive and final [10 10 Grand Slam 10] event of his illustrious career , turned back the clock at Stadium Court with a flowing 6-3 6-3 6-3 serve-and-volley victory over the fifth-seeded [11 11 Dutchman 11] . " It 's a

win that I can be proud of , " said the 30-year-old [12 12 Swede 12] , winner of two [13 13 U.S. 13] Opens and six [14 14 Grand Slam 14] titles in all . " It 's never easy to beat the [15 15 Wimbledon 15] champion . " [16 16 Edberg 16] , who has said he will retire at season 's end , made it look easy under gray skies at the [17 17 National Tennis Centre 17] . The unseeded [18 18 Swede 18] struck quickly , breaking [19 19 Krajicek 19] in the first game and never let loose his grip on the one hour 44 minute match as he served and volleyed with the grace that made him one of the dominant players of his time . " There 's not doubt about it , [20 20 Richard 20] was definitely off his game and I took advantage , " said [21 21 Edberg 21] . " I still have my days where I feel great out there . " Also reaching the second round were top-seeded defending champion [22 22 Pete Sampras 22] , a 6-2 6-2 6-1 winner over last minute replacement [23 23 Jimy Szymanski 23] of [24 24 Venezuela 24] , called on after [25 25 Adrian Voinea 25] of [26 26 Romania 26] withdrew because of a sprained ankle . Third seed [27 27 Thomas Muster 27] of [28 28 Austria 28] also charged into the second round with a 6-1 7-6 (7-2) 6-2 romp over [29 29 Javier Frana 29] of [30 30 Argentina 30] . [31 31 Marcelo Rios 31] of [32 32 Chile 32] , the 10th seed , also advanced . [33 33 Rios 33] claimed a 4-6 6-1 6-4 6-2 victory over [34 34 Romania 34] 's [35 35 Andrei Pavel 35] . On the women 's side , second seed [36 36 Monica Seles 36] got off to a strong start by beating fellow-American Anne Miller 6-0 6-1 and was joined in the second round by [37 37 Spain 37] 's [38 38 Arantxa Sanchez Vicario 38] (seeded third) , 1 Olympic champion [39 39 Lindsay Davenport 39] (8) and [40 40 Karina Habsudova 40] of [41 41 Slovakia 41] (17) . The women 's draw lost another seed when [42 42 Austrian 42] [43 43 Judith Wiesner 43] overcame [44 44 Iva Majoli 44] of [45 45 Croatia 45] 2-6 6-3 6-1 . The fifth- seeded [46 46 Majoli 46] joined [47 47 Anke Huber 47] (5) and [48 48 Magdalena Maleeva 48] (12) on the sidelines .

- 1: 391932 Stefan_Edberg
- 2: 91195 The_Championships,_Wimbledon
- 3: 645042 New_York_City
- 4: 391932 Stefan_Edberg
- 5: 197638 Grand_Slam_(tennis)
- 6: 91195 The_Championships,_Wimbledon
- 7: 161972 Richard_Krajicek
- 8: 82729 US_Open_(tennis)
- 9: 391932 Stefan_Edberg
- 10: 197638 Grand_Slam_(tennis)
- 11: 21148 Netherlands
- 12: 5058739 Sweden
- 13: 3434750 United_States
- 14: 197638 Grand_Slam_(tennis)
- 15: 91195 The_Championships,_Wimbledon
- 16: 391932 Stefan_Edberg
- 17: 3523882 National_Tennis_Centre_(United_Kingdom)

```
18: 5058739 Sweden
19: 161972 Richard_Krajicek
20: 161972 Richard_Krajicek
21: 391932 Stefan_Edberg
22: 65961 Pete_Sampras
23: 3930825 Jimy_Szymanski
24: 32374 Venezuela
25: 4036370 Adrian_Voinea
26: 25445 Romania
27: 166170 Thomas_Muster
28: 26964606 Austria
29: 3930198 Javier_Frana
30: 18951905 Argentina
31: 483554 Marcelo_Ríos
32: 5489 Chile
33: 483554 Marcelo_Ríos
34: 25445 Romania
35: 1723598 Andrei_Pavel
36: 300468 Monica_Seles
37: 26667 Spain
38: 534624 Arantxa_Sánchez_Vicario
39: 52828 Lindsay_Davenport
40: 3062115 Karina_Habšudová
41: 26830 Slovakia
42: 26964606 Austria
43: 2805420 Judith_Wiesner
44: 928220 Iva_Majoli
45: 5573 Croatia
46: 928220 Iva_Majoli
47: 1443478 Anke_Huber
48: 1111175 Magdalena_Maleeva
1: 650360 Olympic_sports
    [winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.25799578, final_local_scores=-0.29054117,
pem=-3.2532303, similarity=13.110958|
     RECALL MISS: 650360 Olympic_sports]
1: 391932 Stefan_Edberg
    [pem_pos=0, final_scores=0.7876802,
final_local_scores=0.79001963, pem=0.0, similarity=27.665474]
2: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.5416696,
final_local_scores=0.51826745, pem=-0.8505555,
similarity=24.181301]
3: 645042 New_York_City
    [pem_pos=1, final_scores=0.36213017,
final_local_scores=0.43985116, pem=-1.7906687,
similarity=26.953684]
4: 391932 Stefan_Edberg
    [pem_pos=0, final_scores=0.748059,
final_local_scores=0.74403805, pem=0.0, similarity=26.293293]
5: 197638 Grand_Slam_(tennis)
    [pem_pos=0, final_scores=0.502395,
```

```
final_local_scores=0.47637773, pem=0.0, similarity=18.305775]
6: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.5151182,
final_local_scores=0.48926407, pem=-0.8505555,
similarity=23.315783]
7: 161972 Richard_Krajicek
    [pem_pos=0, final_scores=0.84851104,
final_local_scores=0.8763189, pem=0.0, similarity=30.240816]
8: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.49732846,
final_local_scores=0.48217928, pem=-1.2370842,
similarity=25.20636]
9: 391932 Stefan_Edberg
    [pem_pos=0, final_scores=0.6306792,
final_local_scores=0.61851984, pem=0.0, similarity=22.547579]
10: 197638 Grand_Slam_(tennis)
    [pem_pos=0, final_scores=0.52955043,
final_local_scores=0.50604093, pem=-0.37449613,
similarity=21.227552]
11: 21148 Netherlands
    [pem_pos=0, final_scores=0.3583355,
final_local_scores=0.40031284, pem=-0.43687075,
similarity=18.411617]
12: 5058739 Sweden
    [pem_pos=0, final_scores=0.4580449,
final_local_scores=0.5100756, pem=-0.6464619,
similarity=22.826946]
13: 3434750 United_States
    [pem_pos=0, final_scores=0.38225442,
final_local_scores=0.4486875, pem=-0.14075686,
similarity=18.2449]
14: 197638 Grand_Slam_(tennis)
    [pem_pos=0, final_scores=0.5489575,
final_local_scores=0.5272403, pem=-0.37449613,
similarity=21.860184]
15: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.54692143,
final_local_scores=0.5240043, pem=-0.8505555,
similarity=24.352499]
16: 391932 Stefan_Edberg
    [pem_pos=0, final_scores=0.7118022,
final_local_scores=0.7071344, pem=0.0, similarity=25.192015]
17: 3523882 National_Tennis_Centre_(United_Kingdom)
    [pem_pos=1, final_scores=0.16941468,
final_local_scores=0.24932238, pem=-0.7567729,
similarity=15.645439]
18: 5058739 Sweden
    [pem_pos=0, final_scores=0.46515656,
final_local_scores=0.517844, pem=-0.6464619,
similarity=23.058771]
19: 161972 Richard_Krajicek
    [pem_pos=0, final_scores=0.35770184,
```

```
final_local_scores=0.34018433, pem=0.0, similarity=14.241492]
20: 161972 Richard_Krajicek
    [pem_pos=0, final_scores=0.94183546,
final_local_scores=0.97826153, pem=0.0, similarity=33.28299]
21: 391932 Stefan_Edberg
    [pem_pos=0, final_scores=0.713968,
final_local_scores=0.70950025, pem=0.0, similarity=25.262615]
22: 65961 Pete_Sampras
    [pem_pos=0, final_scores=0.71455127,
final_local_scores=0.70464605, pem=0.0, similarity=25.117756]
23: 3930825 Jimy_Szymanski
    [pem_pos=0, final_scores=0.6249197,
final_local_scores=0.7493237, pem=0.0, similarity=26.451027]
24: 32374 Venezuela
    [pem_pos=0, final_scores=0.5838801,
final_local_scores=0.6682779, pem=-0.21226841,
similarity=25.186808]
25: 4036370 Adrian_Voinea
    [pem_pos=0, final_scores=0.5617375,
final_local_scores=0.66065675, pem=0.0, similarity=23.805029]
26: 25445 Romania
    [pem_pos=0, final_scores=0.70198804,
final_local_scores=0.7799869, pem=-0.21241479,
similarity=28.521223]
27: 166170 Thomas_Muster
    [pem_pos=0, final_scores=0.9881236,
final_local_scores=1.0178957, pem=0.0, similarity=34.465748]
28: 26964606 Austria
    [pem_pos=0, final_scores=0.6033278,
final_local_scores=0.662528, pem=-0.25262418,
similarity=25.234678]
29: 3930198 Javier_Frana
    [pem_pos=0, final_scores=0.5085408,
final_local_scores=0.5505445, pem=0.0, similarity=20.519058]
30: 18951905 Argentina
    [pem_pos=0, final_scores=0.5520798,
final_local_scores=0.6208612, pem=-0.30003807,
similarity=24.249102]
31: 483554 Marcelo_Rios
    [pem_pos=0, final_scores=0.70663285,
final_local_scores=0.7133846, pem=0.0, similarity=25.378532]
32: 5489 Chile
    [pem_pos=0, final_scores=0.5641689,
final_local_scores=0.66002554, pem=-0.27404457,
similarity=25.276487]
33: 483554 Marcelo_Ríos
    [pem_pos=0, final_scores=0.882016,
final_local_scores=0.90830797, pem=0.0, similarity=31.195435]
34: 25445 Romania
    [pem_pos=0, final_scores=0.63893414,
final_local_scores=0.7111101, pem=-0.21241479,
similarity=26.465801]
```

```
35: 1723598 Andrei_Pavel
    [pem_pos=0, final_scores=0.71530753,
final_local_scores=0.746324, pem=0.0, similarity=26.36151]
36: 300468 Monica_Seles
    [pem_pos=0, final_scores=0.74252814,
final_local_scores=0.7508055, pem=0.0, similarity=26.495247]
37: 26667 Spain
    [pem_pos=0, final_scores=0.62189287,
final_local_scores=0.6986378, pem=-0.18857051,
similarity=25.963934]
38: 534624 Arantxa_Sánchez_Vicario
    [pem_pos=0, final_scores=0.62234265,
final_local_scores=0.6434445, pem=0.0, similarity=23.29138]
39: 52828 Lindsay_Davenport
    [pem_pos=0, final_scores=0.8391703,
final_local_scores=0.8650164, pem=0.0, similarity=29.903526]
40: 3062115 Karina_Habšudová
    [pem_pos=0, final_scores=0.7951028,
final_local_scores=0.9159108, pem=0.0, similarity=31.422318]
41: 26830 Slovakia
    [pem_pos=0, final_scores=0.5830453,
final_local_scores=0.6346258, pem=-0.36602598,
similarity=25.018719]
42: 26964606 Austria
    [pem_pos=0, final_scores=0.6600265,
final_local_scores=0.7244626, pem=-0.40674555,
similarity=27.921068]
43: 2805420 Judith_Wiesner
    [pem_pos=0, final_scores=0.5952515,
final_local_scores=0.69000757, pem=0.0, similarity=24.680916]
44: 928220 Iva_Majoli
    [pem_pos=0, final_scores=0.91101843,
final_local_scores=0.96919554, pem=0.0, similarity=33.01244]
45: 5573 Croatia
    [pem_pos=0, final_scores=0.45573747,
final_local_scores=0.50586265, pem=-0.35021794,
similarity=21.090202]
46: 928220 Iva_Majoli
    [pem_pos=0, final_scores=0.45592934,
final_local_scores=0.47207987, pem=0.0, similarity=18.177519]
47: 1443478 Anke_Huber
    [pem_pos=0, final_scores=0.86629003,
final_local_scores=0.9322625, pem=0.0, similarity=31.910284]
48: 1111175 Magdalena_Maleeva
    [pem_pos=0, final_scores=0.7276059,
final_local_scores=0.7703342, pem=0.0, similarity=27.078022]
```

Sample: b'530_TENNIS'

TENNIS - TUESDAY 'S RESULTS FROM [1 1 U.S. OPEN 1] . [2 2 NEW YORK 2] 1996-08-27 Results of first round matches on Tuesday in the [3 3 U.S. Open 3] tennis championships at the [4 4 National Tennis Centre 4] (prefix denotes seeding) : Women 's

singles 2 - [5 5 Monica Seles 5] ([6 6 U.S. 6]) beat Anne Miller ([7 7 U.S. 7]) 6-0 6-1 Rita Grande ([8 8 Italy 8]) beat [9 9 Alexia Dechaume-Balleret 9] ([10 10 France 10]) 6-3 6-0 [11 11 Judith Wiesner 11] ([12 12 Austria 12]) beat 5 - [13 13 Iva Majoli 13] ([14 14 Croatia 14]) 2-6 6-3 6- 1 Men 's singles 3 - [15 15 Thomas Muster 15] ([16 16 Austria 16]) beat [17 17 Javier Frana 17] ([18 18 Argentina 18]) 6-1 7-6 (7-2) 6-2 Men 's singles 1 - [19 19 Pete Sampras 19] ([20 20 U.S. 20]) beat [21 21 Jimy Szymanski 21] ([22 22 Venezuela 22]) 6-2 6-2 6 - 1 [23 23 Jiri Novak 23] ([24 24 Czech Republic 24]) beat Ben Ellwood ([25 25 Australia 25]) 6-2 6 - 4 6-3 Women 's singles [26 26 Mariaan de Swardt 26] ([27 27 South Africa 27]) beat [28 28 Dominique Van Roost 28] ([29 29 Belgium 29]) 1-6 6-2 7-6 (7-4) [30 30 Florencia La bat 30] ([31 31 Argentina 31]) beat [32 32 Kathy Rinaldi St unkel 32] ([33 33 U.S. 33]) 6 - 2 6-2 [34 34 Nathalie Tauzi at 34] ([35 35 France 35]) beat [36 36 Angelica Gavaldon 36] ([37 37 Mexico 37]) 7-6 (7-4) 6-2 [38 38 Paola Suarez 38] ([39 39 Argentina 39]) beat [40 40 Marianne Werdel Witmeyer 40] ([41 41 U.S. 41]) 6 - 4 6-3 [42 42 Ann Grossman 42] ([43 43 U.S. 43]) beat [44 44 Silvia Farina 44] ([45 45 Italy 45]) 6-4 6-3 Men 's singles [46 46 Alex Corretja 46] ([47 47 Spain 47]) beat [48 48 Byron Black 48] ([49 49 Zimbabwe 49]) 7-6 (8-6) 3-6 6-2 6-2 [50 50 Scott Draper 50] ([51 51 Australia 51]) beat [52 52 Galo Blanco 52] ([53 53 Spain 53]) 6-3 7-5 6-3 [54 54 Petr Korda 54] ([55 55 Czech Republic 55]) beat David Caldwell ([56 56 U.S. 56]) 6-3 3-6 6-3 7-5 [57 57 Bohdan Ulihrach 57] ([58 58 Czech Republic 58]) beat 14 - Alberto Costa ([59 59 Spain 59]) 2-6 6-4 7-6 (7-2) 3-6 6-1 [60 60 Bernd Karbacher 60] ([61 61 Germany 61]) beat [62 62 Jonathan Stark 62] ([63 63 U.S. 63]) 7-5 6-3 5- 7 7-5 Women 's singles 8 - [64 64 Lindsay Davenport 64] ([65 65 U.S. 65]) beat Adriana Serra-Zanetti ([66 66 Italy 66]) 6 - 2 6-1 [67 67 Elena Wagner 67] ([68 68 Germany 68]) beat [69 69 Gigi Fernandez 69] ([70 70 U.S. 70]) 6-1 6-4 [71 71 Kristie Boogert 71] ([72 72 Netherlands 72]) beat [73 73 Joannette Kruger 73] ([74 74 South Africa 74]) 6-1 6-0 Men 's singles [75 75 Stefan Edberg 75] ([76 76 Sweden 76]) beat 5 - [77 77 Richard Krajicek 77] ([78 78 Netherlands 78]) 6- 3 6-3 6-3 10 - [79 79 Marcelo Rios 79] ([80 80 Chile 80]) beat [81 81 Andrei Pavel 81] ([82 82 Romania 82]) 4-6 6-1 6-4 6-2 Women 's singles 3 - [83 83 Arantxa Sanchez Vicario 83] ([84 84 Spain 84]) beat [85 85 Laxmi Poruri 85] ([86 86 U.S. 86]) 6-2 6-1 Men 's singles [87 87 Andrei Olho vskiy 87] ([88 88 Russia 88]) beat [89 89 Pat Cash 89] ([90 90 Australia 90]) 6-4 6-3 6-2 Filippo Veglio ([91 91 Switzerland 91]) beat [92 92 Christian Ruud 92] ([93 93 Norway 93]) 1-6 6 - 2 6-4 6-4 [94 94 Tim Henman 94] ([95 95 Britain 95]) beat Roberto Jabali ([96 96 Brazil 96]) 6-2 6-3 6-4 Pablo Campana ([97 97 Ecuador 97]) beat [98 98 Todd Woodbridge 98] ([99 99 Australia 99]) 6-2 4- 6 6-2 6-4 Herman Gumy ([100 100 Argentina 100]) beat [101 101 Martin Damm 101] ([102 102]

Czech Republic 102]) 7-5 6 - 4 7-5 Jacob Hlasek ([103 103 Switzerland 103]) beat [104 104 Nicklas Kulti 104] ([105 105 Sweden 105]) 6-3 6-4 4-6 6-4 Women 's singles 17- [106 106 Karina Habsudova 106] ([107 107 Slovakia 107]) beat Radka Bobkova ([108 108 Czech Republic 108]) 6-4 6-1 [109 109 Karin Kschwendt 109] ([110 110 Austria 110]) beat [111 111 Sandra Kleinova 111] ([112 112 Czech Republic 112]) 6-3 6-4 Annabel Ellwood ([113 113 Australia 113]) beat [114 114 Jennifer Capriati 114] ([115 115 U.S. 115]) 6-4 6 - 4 [116 116 Nicole Arendt 116] ([117 117 U.S. 117]) beat [118 118 Sandra Cacic 118] ([119 119 U.S. 119]) 6-2 7-6 (8-6) [120 120 Elena Likhovtseva 120] ([121 121 Russia 121]) beat [122 122 Kyoko Nagatsuka 122] ([123 123 Japan 123]) 7-6 (7-5) 6-1 [124 124 Sandrine Testud 124] ([125 125 France 125]) beat [126 126 Pam Shriver 126] ([127 127 U.S. 127]) 7-5 6-2 Women 's singles [128 128 Kimberly Po 128] ([129 129 U.S. 129]) beat 10 - [130 130 Kimiko Date 130] ([131 131 Japan 131]) 6-2 7-5 [132 132 Natasha Zvereva 132] ([133 133 Belarus 133]) beat [134 134 Virginia Ruano-Pascual 134] ([135 135 Spain 135]) 6 - 2 6-7 (5-7) 6-2 Tina Kirzan ([136 136 Slovakia 136]) beat [137 137 Rika Hiraki 137] ([138 138 Japan 138]) 7-6 (7-4) 7-5 [139 139 Petra Langrova 139] ([140 140 Czech Republic 140]) beat Karina Adams ([141 141 U.S. 141]) 6-4 6-2 [142 142 Tami Whitlinger Jones 142] ([143 143 U.S. 143]) beat [144 144 Sandra Cecchini 144] ([145 145 Italy 145]) 6-2 6-0 7 - [146 146 Jana Novotna 146] ([147 147 Czech Republic 147]) beat [148 148 Francesca Lubiani 148] ([149 149 Italy 149]) 6-1 7-5 Men 's singles 13 - [150 150 Thomas Enqvist 150] ([151 151 Sweden 151]) beat [152 152 Stephane Simian 152] ([153 153 France 153]) 6-3 6-1 6-4 Men 's singles [154 154 Mikael Tillstrom 154] ([155 155 Sweden 155]) beat Tamer El Sawy ([156 156 Egypt 156]) 1-6 7-6 (9 - 7) 6-1 3-6 6-4 [157 157 Roberto Carretero 157] ([158 158 Spain 158]) beat [159 159 Jordi Burillo 159] ([160 160 Spain 160]) 6-3 4-6 6-0 1-0 Retired (ankle injury) [161 161 Thomas Johansson 161] ([162 162 Sweden 162]) beat [163 163 Renzo Furlan 163] ([164 164 Italy 164]) 4-6 2-6 7-5 6-1 7-5 [165 165 Mark Knowles 165] ([166 166 Bahamas 166]) beat [167 167 Marcelo Filippini 167] ([168 168 Uruguay 168]) 6-3 7-5 6-1 [169 169 Jared Palmer 169] ([170 170 U.S. 170]) beat 15 - [171 171 Marc Rosset 171] ([172 172 Switzerland 172]) 6-7 (7-9) 6-4 6-4 6-3 Women 's singles [173 173 Amy Frazier 173] ([174 174 U.S. 174]) beat [175 175 Larisa Neiland 175] ([176 176 Latvia 176]) 6-1 6-3 Women 's singles [177 177 Lisa Raymond 177] ([178 178 U.S. 178]) beat [179 179 Lori McNeil 179] ([180 180 U.S. 180]) 7-6 (8-6) 6-3 Sandra Dopfer ([181 181 Austria 181]) beat [182 182 Zina Garrison Jackson 182] ([183 183 U.S. 183]) 2-6 6-3 7-5 4 - [184 184 Conchita Martinez 184] ([185 185 Spain 185]) beat [186 186 Ruxandra Dragomir 186] ([187 187 Romania 187]) 6-2 6-0 [188 188 Naoko Sawamatsu 188] ([189 189 Japan 189]) beat [190 190 Rennae Stubbs 190] ([191 191 Australia 191]) 6-4 6-3 [192 192 Miriam Oremans 192] ([193 192])

3 Netherlands 193]) beat [194 194 Radka Zrubakova 194] ([195
195 Slovakia 195]) 6-2 4-6 6-1 Men 's singles [196 196 Doug F
lach 196] ([197 197 U.S. 197]) beat [198 198 Gianluca Pozzi
198] ([199 199 Italy 199]) 7-5 7-6 (7-5) 2-6 7-6 (8-6)
16 - [200 200 Cedric Pioline 200] ([201 201 France 201])
beat [202 202 Francisco Clavet 202] ([203 203 Spain 203])
6-4 7-6 (7-3) 6-4 [204 204 Javier Sanchez 204] ([205 205 Sp
ain 205]) beat [206 206 David Skoch 206] ([207 207 Czech Re
public 207]) 6-2 7-6 (7-0) 6-3 Women 's singles , first
round 1 - [208 208 Steffi Graf 208] ([209 209 Germany 209])
beat [210 210 Yayuk Basuki 210] ([211 211 Indonesia 211])
6-3 7-6 (7-4)

1: 82729 US_Open_(tennis)
2: 645042 New_York_City
3: 82729 US_Open_(tennis)
4: 3523882 National_Tennis_Centre_(United_Kingdom)
5: 300468 Monica_Seles
6: 3434750 United_States
7: 3434750 United_States
8: 14532 Italy
9: 16570128 Alexia_Dechaume-Balleret
10: 5843419 France
11: 2805420 Judith_Wiesner
12: 26964606 Austria
13: 928220 Iva_Majoli
14: 5573 Croatia
15: 166170 Thomas_Muster
16: 26964606 Austria
17: 3930198 Javier_Frana
18: 18951905 Argentina
19: 65961 Pete_Sampras
20: 3434750 United_States
21: 3930825 Jimy_Szymanski
22: 32374 Venezuela
23: 1554049 Jiří_Novák
24: 5321 Czech_Republic
25: 4689264 Australia
26: 10119577 Mariaan_de_Swardt
27: 17416221 South_Africa
28: 1849753 Dominique_Monami
29: 3343 Belgium
30: 3128595 Florencia_Labat
31: 18951905 Argentina
32: 4077090 Kathy_Rinaldi
33: 3434750 United_States
34: 1028366 Nathalie_Tauziat
35: 5843419 France
36: 9789214 Angélica_Gavaldón
37: 3966054 Mexico
38: 705473 Paola_Suárez
39: 18951905 Argentina
40: 15113825 Marianne_Werdel

41: 3434750 United_States
42: 15097221 Ann_Grossman
43: 3434750 United_States
44: 524346 Silvia_Farina_Elia
45: 14532 Italy
46: 17733765 Alex_Corretja
47: 26667 Spain
48: 1869714 Byron_Black
49: 34399 Zimbabwe
50: 3983307 Scott_Draper
51: 4689264 Australia
52: 441573 Galo_Blanco
53: 26667 Spain
54: 1439769 Petr_Korda
55: 5321 Czech_Republic
56: 3434750 United_States
57: 2805841 Bohdan_Ulihrach
58: 5321 Czech_Republic
59: 26667 Spain
60: 15290626 Bernd_Karbacher
61: 11867 Germany
62: 3079144 Jonathan_Stark_(tennis)
63: 3434750 United_States
64: 52828 Lindsay_Davenport
65: 3434750 United_States
66: 14532 Italy
67: 18330817 Elena_Pampoulova
68: 11867 Germany
69: 237455 Gigi_Fernández
70: 3434750 United_States
71: 2527532 Kristie_Boogert
72: 21148 Netherlands
73: 12695931 Joannette_Kruger
74: 17416221 South_Africa
75: 391932 Stefan_Edberg
76: 5058739 Sweden
77: 161972 Richard_Krajicek
78: 21148 Netherlands
79: 483554 Marcelo_Ríos
80: 5489 Chile
81: 1723598 Andrei_Pavel
82: 25445 Romania
83: 534624 Arantxa_Sánchez_Vicario
84: 26667 Spain
85: 10679190 Laxmi_Poruri
86: 3434750 United_States
87: 3930485 Andrei_Olhovskiy
88: 25391 Russia
89: 435558 Pat_Cash
90: 4689264 Australia
91: 26748 Switzerland
92: 3929692 Christian_Ruud

93: 21241 Norway
94: 262030 Tim_Henman
95: 31717 United_Kingdom
96: 3383 Brazil
97: 9334 Ecuador
98: 1547951 Todd_Woodbridge
99: 4689264 Australia
100: 18951905 Argentina
101: 4026614 Martin_Damm
102: 5321 Czech_Republic
103: 26748 Switzerland
104: 2987792 Nicklas_Kulti
105: 5058739 Sweden
106: 3062115 Karina_Habšudová
107: 26830 Slovakia
108: 5321 Czech_Republic
109: 27388154 Karin_Kschwendt
110: 26964606 Austria
111: 14851318 Sandra_Kleinová
112: 5321 Czech_Republic
113: 4689264 Australia
114: 325049 Jennifer_Capriati
115: 3434750 United_States
116: 4038024 Nicole_Arendt
117: 3434750 United_States
118: 15290311 Sandra_Cacic
119: 3434750 United_States
120: 2350072 Elena_Likhovtseva
121: 25391 Russia
122: 27399536 Kyōko_Nagatsuka
123: 15573 Japan
124: 3127711 Sandrine_Testud
125: 5843419 France
126: 1399994 Pam_Shriver
127: 3434750 United_States
128: 6426517 Kimberly_Po
129: 3434750 United_States
130: 495296 Kimiko_Date-Krumm
131: 15573 Japan
132: 916019 Natasha_Zvereva
133: 3457 Belarus
134: 1858321 Virginia_Ruano_Pascual
135: 26667 Spain
136: 26830 Slovakia
137: 5504097 Rika_Hiraki
138: 15573 Japan
139: 19609793 Petra_Langrová
140: 5321 Czech_Republic
141: 3434750 United_States
142: 2783800 Tami_Whitlinger
143: 3434750 United_States
144: 14979843 Sandra_Cecchini

145: 14532 Italy
146: 307050 Jana_Novotná
147: 5321 Czech_Republic
148: 26914284 Francesca_Lubiani
149: 14532 Italy
150: 933709 Thomas_Enqvist
151: 5058739 Sweden
152: 22577800 Stéphane_Simian
153: 5843419 France
154: 3943969 Mikael_Tillström
155: 5058739 Sweden
156: 8087628 Egypt
157: 1723684 Roberto_Carretero
158: 26667 Spain
159: 14891740 Jordi_Burillo
160: 26667 Spain
161: 1408278 Thomas_Johansson
162: 5058739 Sweden
163: 3929800 Renzo_Furlan
164: 14532 Italy
165: 2261491 Mark_Knowles
166: 3451 The_Bahamas
167: 2936118 Marcelo_Filippini
168: 31843 Uruguay
169: 3056059 Jared_Palmer
170: 3434750 United_States
171: 1717291 Marc_Rosset
172: 26748 Switzerland
173: 2634301 Amy_Frazier
174: 3434750 United_States
175: 3877687 Larisa_Neiland
176: 17514 Latvia
177: 2199290 Lisa_Raymond
178: 3434750 United_States
179: 6426781 Lori_McNeil
180: 3434750 United_States
181: 26964606 Austria
182: 1028356 Zina_Garrison
183: 3434750 United_States
184: 784152 Conchitá_Martínez
185: 26667 Spain
186: 4038121 Ruxandra_Dragomir
187: 25445 Romania
188: 2631927 Naoko_Sawamatsu
189: 15573 Japan
190: 2482750 Rennae_Stubbs
191: 4689264 Australia
192: 2527510 Miriam_Oremans
193: 21148 Netherlands
194: 15443594 Radka_Zrubáková
195: 26830 Slovakia
196: 4025956 Doug_Flach

```
197: 3434750 United_States
198: 4026677 Gianluca_Pozzi
199: 14532 Italy
200: 979076 Cédric_Pioline
201: 5843419 France
202: 4076952 Francisco_Clavet
203: 26667 Spain
204: 420527 Javier_Sánchez
205: 26667 Spain
206: 12871566 David_Škoch
207: 5321 Czech_Republic
208: 164640 Steffi_Graf
209: 11867 Germany
210: 2672548 Yayuk_Basuki
211: 14579 Indonesia
1: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.5190733,
final_local_scores=0.53712595, pem=-1.2370842,
similarity=26.846079]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.42493635,
final_local_scores=0.45673782, pem=-1.7906687,
similarity=27.457615]
3: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.5123475,
final_local_scores=0.529779, pem=-1.2370842,
similarity=26.626831]
4: 3523882 National_Tennis_Centre_(United_Kingdom)
    [pem_pos=1, final_scores=0.045273706,
final_local_scores=0.11216408, pem=-0.7567729,
similarity=11.5523615]
5: 300468 Monica_Seles
    [pem_pos=0, final_scores=0.82029754,
final_local_scores=0.87602586, pem=0.0, similarity=30.232073]
6: 3434750 United_States
    [pem_pos=0, final_scores=0.48529774,
final_local_scores=0.4825409, pem=-0.14075686,
similarity=19.255154]
7: 3434750 United_States
    [pem_pos=0, final_scores=0.51597685,
final_local_scores=0.51605326, pem=-0.14075686,
similarity=20.255228]
8: 14532 Italy
    [pem_pos=0, final_scores=0.75318414,
final_local_scores=0.79431707, pem=-0.20612453,
similarity=28.914656]
9: 16570128 Alexia_Dechaume-Balleret
    [pem_pos=0, final_scores=0.7932671,
final_local_scores=0.93342894, pem=0.0, similarity=31.945095]
10: 5843419 France
    [pem_pos=0, final_scores=0.76348823,
final_local_scores=0.7780127, pem=-0.16240834,
```

```
similarity=28.190365]
11: 2805420 Judith_Wiesner
    [pem_pos=0, final_scores=0.584921,
final_local_scores=0.69855756, pem=0.0, similarity=24.936064]
12: 26964606 Austria
    [pem_pos=0, final_scores=0.60446984,
final_local_scores=0.60426736, pem=-0.25262418,
similarity=23.496065]
13: 928220 Iva_Majoli
    [pem_pos=0, final_scores=0.87368983,
final_local_scores=0.96524566, pem=0.0, similarity=32.89457]
14: 5573 Croatia
    [pem_pos=0, final_scores=0.6217216,
final_local_scores=0.6329655, pem=-0.35021794,
similarity=24.883205]
15: 166170 Thomas_Muster
    [pem_pos=0, final_scores=0.8874213,
final_local_scores=0.9571863, pem=0.0, similarity=32.65406]
16: 26964606 Austria
    [pem_pos=0, final_scores=0.641079,
final_local_scores=0.64425737, pem=-0.25262418,
similarity=24.689447]
17: 3930198 Javier_Frana
    [pem_pos=0, final_scores=0.47785378,
final_local_scores=0.5479559, pem=0.0, similarity=20.441809]
18: 18951905 Argentina
    [pem_pos=0, final_scores=0.60663646,
final_local_scores=0.6133088, pem=-0.30003807,
similarity=24.023724]
19: 65961 Pete_Sampras
    [pem_pos=0, final_scores=0.700585,
final_local_scores=0.7418889, pem=0.0, similarity=26.229156]
20: 3434750 United_States
    [pem_pos=0, final_scores=0.5008117,
final_local_scores=0.49948758, pem=-0.14075686,
similarity=19.760878]
21: 3930825 Jimy_Szymanski
    [pem_pos=0, final_scores=0.56961817,
final_local_scores=0.69555193, pem=0.0, similarity=24.84637]
22: 32374 Venezuela
    [pem_pos=0, final_scores=0.58891267,
final_local_scores=0.6112575, pem=-0.21226841,
similarity=23.485205]
23: 1554049 Jiří_Novák
    [pem_pos=0, final_scores=0.81236786,
final_local_scores=0.9150222, pem=0.0, similarity=31.3958]
24: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.6252356,
final_local_scores=0.62825644, pem=-0.27794555,
similarity=24.34965]
25: 4689264 Australia
    [pem_pos=0, final_scores=0.65127575,
```

```
final_local_scores=0.6824804, pem=-0.17168477,  
similarity=25.389938]  
26: 10119577 Mariaan_de_Swardt  
    [pem_pos=0, final_scores=0.6586009,  
final_local_scores=0.7355042, pem=0.0, similarity=26.038626]  
27: 17416221 South_Africa  
    [pem_pos=0, final_scores=0.6497794,  
final_local_scores=0.70675087, pem=-0.26251382,  
similarity=26.60816]  
28: 1849753 Dominique_Monami  
    [pem_pos=0, final_scores=0.67122275,  
final_local_scores=0.78297716, pem=0.0, similarity=27.455313]  
29: 3343 Belgium  
    [pem_pos=0, final_scores=0.7287685,  
final_local_scores=0.7323401, pem=-0.24180195,  
similarity=27.25916]  
30: 3128595 Florencia_Labat  
    [pem_pos=0, final_scores=0.54507023,  
final_local_scores=0.65634227, pem=0.0, similarity=23.676275]  
31: 18951905 Argentina  
    [pem_pos=0, final_scores=0.6976469,  
final_local_scores=0.7127239, pem=-0.30003807,  
similarity=26.990469]  
32: 4077090 Kathy_Rinaldi  
    [pem_pos=0, final_scores=0.5144382,  
final_local_scores=0.62510705, pem=0.0, similarity=22.744154]  
33: 3434750 United_States  
    [pem_pos=0, final_scores=0.53254235,  
final_local_scores=0.5341485, pem=-0.14075686,  
similarity=20.795229]  
34: 1028366 Nathalie_Tauziat  
    [pem_pos=0, final_scores=0.7438044,  
final_local_scores=0.85330397, pem=0.0, similarity=29.554005]  
35: 5843419 France  
    [pem_pos=0, final_scores=0.73335403,  
final_local_scores=0.7450956, pem=-0.16240834,  
similarity=27.208054]  
36: 9789214 Angélica_Gavaldón  
    [pem_pos=0, final_scores=0.4187039,  
final_local_scores=0.513963, pem=0.0, similarity=19.427393]  
37: 3966054 Mexico  
    [pem_pos=0, final_scores=0.66239816,  
final_local_scores=0.68717206, pem=-0.18258883,  
similarity=25.589245]  
38: 705473 Paola_Suárez  
    [pem_pos=0, final_scores=0.70885336,  
final_local_scores=0.7792573, pem=0.0, similarity=27.344305]  
39: 18951905 Argentina  
    [pem_pos=0, final_scores=0.69812137,  
final_local_scores=0.7132422, pem=-0.30003807,  
similarity=27.005936]  
40: 15113825 Marianne_Werdel
```

```
[pem_pos=0, final_scores=0.3865987,
final_local_scores=0.50233656, pem=0.0, similarity=19.080439]
41: 3434750 United_States
    [pem_pos=0, final_scores=0.527564,
final_local_scores=0.5287104, pem=-0.14075686,
similarity=20.632944]
42: 15097221 Ann_Grossman
    [pem_pos=0, final_scores=0.8259736,
final_local_scores=0.9683396, pem=0.0, similarity=32.9869]
43: 3434750 United_States
    [pem_pos=0, final_scores=0.51794297,
final_local_scores=0.51820093, pem=-0.14075686,
similarity=20.31932]
44: 524346 Silvia_Farina_Elia
    [pem_pos=0, final_scores=0.5547918,
final_local_scores=0.63662076, pem=0.0, similarity=23.087746]
45: 14532 Italy
    [pem_pos=0, final_scores=0.72106093,
final_local_scores=0.75922734, pem=-0.20612453,
similarity=27.867508]
46: 17733765 Àlex_Corretja
    [pem_pos=0, final_scores=0.86312777,
final_local_scores=0.9365644, pem=0.0, similarity=32.038662]
47: 26667 Spain
    [pem_pos=0, final_scores=0.652737,
final_local_scores=0.66237813, pem=-0.18857051,
similarity=24.881874]
48: 1869714 Byron_Black
    [pem_pos=0, final_scores=0.646768,
final_local_scores=0.747895, pem=0.0, similarity=26.408392]
49: 34399 Zimbabwe
    [pem_pos=0, final_scores=0.50559914,
final_local_scores=0.5703196, pem=-0.36737105,
similarity=23.107008]
50: 3983307 Scott_Draper
    [pem_pos=0, final_scores=0.68001896,
final_local_scores=0.7961345, pem=0.0, similarity=27.847954]
51: 4689264 Australia
    [pem_pos=0, final_scores=0.6969064,
final_local_scores=0.7323249, pem=-0.17168477,
similarity=26.877398]
52: 441573 Galo_Blanco
    [pem_pos=0, final_scores=0.60640275,
final_local_scores=0.71865624, pem=0.0, similarity=25.535849]
53: 26667 Spain
    [pem_pos=0, final_scores=0.6411441,
final_local_scores=0.64971465, pem=-0.18857051,
similarity=24.50397]
54: 1439769 Petr_Korda
    [pem_pos=0, final_scores=0.8551355,
final_local_scores=0.9560723, pem=0.0, similarity=32.620815]
55: 5321 Czech_Republic
```

[pem_pos=0, final_scores=0.59635884,
final_local_scores=0.59671295, pem=-0.27794555,
similarity=23.408329]
56: 3434750 United_States
[pem_pos=0, final_scores=0.5202558,
final_local_scores=0.52072734, pem=-0.14075686,
similarity=20.394714]
57: 2805841 Bohdan_Ulihrach
[pem_pos=0, final_scores=0.77665263,
final_local_scores=0.9017183, pem=0.0, similarity=30.998785]
58: 5321 Czech_Republic
[pem_pos=0, final_scores=0.6297303,
final_local_scores=0.6331662, pem=-0.27794555,
similarity=24.496166]
59: 26667 Spain
[pem_pos=0, final_scores=0.6544938,
final_local_scores=0.66429716, pem=-0.18857051,
similarity=24.93914]
60: 15290626 Bernd_Karbacher
[pem_pos=0, final_scores=0.60675526,
final_local_scores=0.6999246, pem=0.0, similarity=24.976858]
61: 11867 Germany
[pem_pos=0, final_scores=0.9386526,
final_local_scores=0.97442836, pem=-0.15529114,
similarity=34.013096]
62: 3079144 Jonathan_Stark_(tennis)
[pem_pos=1, final_scores=0.67176515,
final_local_scores=0.76617175, pem=-0.86421156,
similarity=31.653524]
63: 3434750 United_States
[pem_pos=0, final_scores=0.52186257,
final_local_scores=0.52248245, pem=-0.14075686,
similarity=20.44709]
64: 52828 Lindsay_Davenport
[pem_pos=0, final_scores=0.84103566,
final_local_scores=0.9133894, pem=0.0, similarity=31.347075]
65: 3434750 United_States
[pem_pos=0, final_scores=0.5005762,
final_local_scores=0.49923033, pem=-0.14075686,
similarity=19.7532]
66: 14532 Italy
[pem_pos=0, final_scores=0.72933716,
final_local_scores=0.7682679, pem=-0.20612453,
similarity=28.137299]
67: 18330817 Elena_Pampoulova
[pem_pos=0, final_scores=0.6431034,
final_local_scores=0.72092706, pem=0.0, similarity=25.603615]
68: 11867 Germany
[pem_pos=0, final_scores=0.8685505,
final_local_scores=0.8978525, pem=-0.15529114,
similarity=31.727919]
69: 237455 Gigi_Fernández

[pem_pos=0, final_scores=0.8471255,
final_local_scores=0.9314725, pem=0.0, similarity=31.88671]
70: 3434750 United_States
[pem_pos=0, final_scores=0.5319827,
final_local_scores=0.5335372, pem=-0.14075686,
similarity=20.776985]
71: 2527532 Kristie_Boogert
[pem_pos=0, final_scores=0.557831,
final_local_scores=0.6866461, pem=0.0, similarity=24.580603]
72: 21148 Netherlands
[pem_pos=0, final_scores=0.60014284,
final_local_scores=0.6063817, pem=-0.24940795,
similarity=23.541672]
73: 12695931 Joannette_Kruger
[pem_pos=0, final_scores=0.6727055,
final_local_scores=0.79527897, pem=0.0, similarity=27.822424]
74: 17416221 South_Africa
[pem_pos=0, final_scores=0.55942446,
final_local_scores=0.60805184, pem=-0.26251382,
similarity=23.662785]
75: 391932 Stefan_Edberg
[pem_pos=0, final_scores=0.81854695,
final_local_scores=0.88861376, pem=0.0, similarity=30.60772]
76: 5058739 Sweden
[pem_pos=0, final_scores=0.780526,
final_local_scores=0.8072409, pem=-0.2135508,
similarity=29.340714]
77: 161972 Richard_Krajicek
[pem_pos=0, final_scores=0.80853087,
final_local_scores=0.89233893, pem=0.0, similarity=30.718887]
78: 21148 Netherlands
[pem_pos=0, final_scores=0.59513146,
final_local_scores=0.60090756, pem=-0.24940795,
similarity=23.378313]
79: 483554 Marcelo_Ríos
[pem_pos=0, final_scores=0.6460507,
final_local_scores=0.69832236, pem=0.0, similarity=24.929045]
80: 5489 Chile
[pem_pos=0, final_scores=0.678684,
final_local_scores=0.7172449, pem=-0.27404457,
similarity=26.984028]
81: 1723598 Andrei_Pavel
[pem_pos=0, final_scores=0.6321708,
final_local_scores=0.7004723, pem=0.0, similarity=24.993204]
82: 25445 Romania
[pem_pos=0, final_scores=0.7212223,
final_local_scores=0.7414183, pem=-0.21241479,
similarity=27.370258]
83: 534624 Arantxa_Sánchez_Vicario
[pem_pos=0, final_scores=0.6343357,
final_local_scores=0.69628215, pem=0.0, similarity=24.86816]
84: 26667 Spain

[pem_pos=0, final_scores=0.67577416,
final_local_scores=0.68754274, pem=-0.18857051,
similarity=25.632835]
85: 10679190 Laxmi_Poruri
[pem_pos=0, final_scores=0.48915702,
final_local_scores=0.6053499, pem=0.0, similarity=22.154562]
86: 3434750 United_States
[pem_pos=0, final_scores=0.53395313,
final_local_scores=0.53568953, pem=-0.14075686,
similarity=20.841213]
87: 3930485 Andrei_Olhovskiy
[pem_pos=0, final_scores=0.81017584,
final_local_scores=0.9256733, pem=0.0, similarity=31.713652]
88: 25391 Russia
[pem_pos=0, final_scores=0.87438184,
final_local_scores=0.93444604, pem=-0.20299491,
similarity=33.079365]
89: 435558 Pat_Cash
[pem_pos=0, final_scores=0.51360196,
final_local_scores=0.56435305, pem=0.0, similarity=20.931133]
90: 4689264 Australia
[pem_pos=0, final_scores=0.67584735,
final_local_scores=0.7093211, pem=-0.17168477,
similarity=26.190918]
91: 26748 Switzerland
[pem_pos=0, final_scores=0.7214659,
final_local_scores=0.7327473, pem=-0.18598008,
similarity=26.967743]
92: 3929692 Christian_Ruud
[pem_pos=0, final_scores=0.5512532,
final_local_scores=0.64979094, pem=0.0, similarity=23.48077]
93: 21241 Norway
[pem_pos=0, final_scores=0.7569789,
final_local_scores=0.81466866, pem=-0.22668563,
similarity=29.633802]
94: 262030 Tim_Henman
[pem_pos=0, final_scores=0.743923,
final_local_scores=0.7911575, pem=0.0, similarity=27.69943]
95: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5421886,
final_local_scores=0.5667415, pem=-0.61114657,
similarity=24.32592]
96: 3383 Brazil
[pem_pos=0, final_scores=0.7928358,
final_local_scores=0.8275099, pem=-0.21313342,
similarity=29.943312]
97: 9334 Ecuador
[pem_pos=0, final_scores=0.49493724,
final_local_scores=0.55397016, pem=-0.21404564,
similarity=21.785301]
98: 1547951 Todd_Woodbridge
[pem_pos=0, final_scores=0.7503796,

```
final_local_scores=0.83959204, pem=0.0, similarity=29.144814]
99: 4689264 Australia
    [pem_pos=0, final_scores=0.66652113,
final_local_scores=0.69913363, pem=-0.17168477,
similarity=25.886904]
100: 18951905 Argentina
    [pem_pos=0, final_scores=0.68184465,
final_local_scores=0.6954623, pem=-0.30003807,
similarity=26.475348]
101: 4026614 Martin_Damm
    [pem_pos=0, final_scores=0.84934586,
final_local_scores=0.9554524, pem=0.0, similarity=32.602318]
102: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.5565577,
final_local_scores=0.5532363, pem=-0.27794555,
similarity=22.110899]
103: 26748 Switzerland
    [pem_pos=0, final_scores=0.6989334,
final_local_scores=0.70813406, pem=-0.18598008,
similarity=26.233234]
104: 2987792 Nicklas_Kulti
    [pem_pos=0, final_scores=0.74413294,
final_local_scores=0.8563644, pem=0.0, similarity=29.645336]
105: 5058739 Sweden
    [pem_pos=0, final_scores=0.758429,
final_local_scores=0.78310335, pem=-0.2135508,
similarity=28.620401]
106: 3062115 Karina_Habšudová
    [pem_pos=0, final_scores=0.7148965,
final_local_scores=0.86792356, pem=0.0, similarity=29.990282]
107: 26830 Slovakia
    [pem_pos=0, final_scores=0.6898621,
final_local_scores=0.6977334, pem=-0.36602598,
similarity=26.901976]
108: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.5901277,
final_local_scores=0.5899064, pem=-0.27794555,
similarity=23.205208]
109: 27388154 Karin_Kschwendt
    [pem_pos=0, final_scores=0.7884039,
final_local_scores=0.90026885, pem=0.0, similarity=30.955532]
110: 26964606 Austria
    [pem_pos=0, final_scores=0.64320284,
final_local_scores=0.6465773, pem=-0.25262418,
similarity=24.758678]
111: 14851318 Sandra_Kleinová
    [pem_pos=0, final_scores=0.620498,
final_local_scores=0.724221, pem=0.0, similarity=25.701912]
112: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.60648793,
final_local_scores=0.6077775, pem=-0.27794555,
similarity=23.738516]
```

113: 4689264 Australia
[pem_pos=0, final_scores=0.7226203,
final_local_scores=0.76041347, pem=-0.17168477,
similarity=27.715616]
114: 325049 Jennifer_Capriati
[pem_pos=0, final_scores=0.81870896,
final_local_scores=0.8787133, pem=0.0, similarity=30.312271]
115: 3434750 United_States
[pem_pos=0, final_scores=0.51093084,
final_local_scores=0.51054126, pem=-0.14075686,
similarity=20.09074]
116: 4038024 Nicole_Arendt
[pem_pos=0, final_scores=0.72464997,
final_local_scores=0.8495062, pem=0.0, similarity=29.440672]
117: 3434750 United_States
[pem_pos=0, final_scores=0.49085444,
final_local_scores=0.4886108, pem=-0.14075686,
similarity=19.436293]
118: 15290311 Sandra_Cacic
[pem_pos=0, final_scores=0.53306425,
final_local_scores=0.6376526, pem=0.0, similarity=23.118538]
119: 3434750 United_States
[pem_pos=0, final_scores=0.5197368,
final_local_scores=0.5201604, pem=-0.14075686,
similarity=20.377794]
120: 2350072 Elena_Likhovtseva
[pem_pos=0, final_scores=0.9018524,
final_local_scores=1.0173929, pem=0.0, similarity=34.45075]
121: 25391 Russia
[pem_pos=0, final_scores=0.81736,
final_local_scores=0.87215847, pem=-0.20299491,
similarity=31.22058]
122: 27399536 Kyōko_Nagatsuka
[pem_pos=0, final_scores=0.65032595,
final_local_scores=0.78624433, pem=0.0, similarity=27.552813]
123: 15573 Japan
[pem_pos=0, final_scores=0.70156115,
final_local_scores=0.74885035, pem=-0.13549899,
similarity=27.173765]
124: 3127711 Sandrine_Testud
[pem_pos=0, final_scores=0.7682851,
final_local_scores=0.8720606, pem=0.0, similarity=30.113739]
125: 5843419 France
[pem_pos=0, final_scores=0.7452846,
final_local_scores=0.758128, pem=-0.16240834,
similarity=27.596966]
126: 1399994 Pam_Shriver
[pem_pos=0, final_scores=0.8817938,
final_local_scores=0.980519, pem=0.0, similarity=33.350357]
127: 3434750 United_States
[pem_pos=0, final_scores=0.518387,
final_local_scores=0.518686, pem=-0.14075686,

```
similarity=20.333796]
128: 6426517 Kimberly_Po
    [pem_pos=0, final_scores=0.6212437,
final_local_scores=0.7307803, pem=0.0, similarity=25.897655]
129: 3434750 United_States
    [pem_pos=0, final_scores=0.4864679,
final_local_scores=0.4838192, pem=-0.14075686,
similarity=19.293299]
130: 495296 Kimiko_Date-Krumm
    [pem_pos=0, final_scores=0.7344965,
final_local_scores=0.817542, pem=0.0, similarity=28.486797]
131: 15573 Japan
    [pem_pos=0, final_scores=0.66474867,
final_local_scores=0.7086383, pem=-0.13549899,
similarity=25.973757]
132: 916019 Natasha_Zvereva
    [pem_pos=0, final_scores=0.87316996,
final_local_scores=0.9807386, pem=0.0, similarity=33.356907]
133: 3457 Belarus
    [pem_pos=0, final_scores=0.59338903,
final_local_scores=0.6108039, pem=-0.4220667,
similarity=24.612583]
134: 1858321 Virginia_Ruano_Pascual
    [pem_pos=0, final_scores=0.7736197,
final_local_scores=0.8614671, pem=0.0, similarity=29.79761]
135: 26667 Spain
    [pem_pos=0, final_scores=0.60570604,
final_local_scores=0.61100394, pem=-0.18857051,
similarity=23.348764]
136: 26830 Slovakia
    [pem_pos=0, final_scores=0.6560318,
final_local_scores=0.66077894, pem=-0.36602598,
similarity=25.799181]
137: 5504097 Rika_Hiraki
    [pem_pos=0, final_scores=0.47998685,
final_local_scores=0.60136414, pem=0.0, similarity=22.035618]
138: 15573 Japan
    [pem_pos=0, final_scores=0.78166395,
final_local_scores=0.83635056, pem=-0.13549899,
similarity=29.784946]
139: 19609793 Petra_Langrová
    [pem_pos=0, final_scores=0.62080187,
final_local_scores=0.73877555, pem=0.0, similarity=26.13625]
140: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.6055014,
final_local_scores=0.6066998, pem=-0.27794555,
similarity=23.706358]
141: 3434750 United_States
    [pem_pos=0, final_scores=0.51551753,
final_local_scores=0.5155515, pem=-0.14075686,
similarity=20.240257]
142: 2783800 Tami_Whitlinger
```

[pem_pos=0, final_scores=0.60763055,
final_local_scores=0.7187854, pem=0.0, similarity=25.539703]
143: 3434750 United_States
[pem_pos=0, final_scores=0.5037597,
final_local_scores=0.5027078, pem=-0.14075686,
similarity=19.856974]
144: 14979843 Sandra_Cecchini
[pem_pos=0, final_scores=0.74575216,
final_local_scores=0.8637652, pem=0.0, similarity=29.866188]
145: 14532 Italy
[pem_pos=0, final_scores=0.7379355,
final_local_scores=0.7776603, pem=-0.20612453,
similarity=28.417585]
146: 307050 Jana_Novotná
[pem_pos=0, final_scores=0.89698595,
final_local_scores=0.9935233, pem=0.0, similarity=33.73843]
147: 5321 Czech_Republic
[pem_pos=0, final_scores=0.5833083,
final_local_scores=0.58245724, pem=-0.27794555,
similarity=22.98291]
148: 26914284 Francesca_Lubiani
[pem_pos=0, final_scores=0.53255314,
final_local_scores=0.66581225, pem=0.0, similarity=23.95888]
149: 14532 Italy
[pem_pos=0, final_scores=0.7578526,
final_local_scores=0.79941666, pem=-0.20612453,
similarity=29.06684]
150: 933709 Thomas_Enqvist
[pem_pos=0, final_scores=0.75893635,
final_local_scores=0.8386889, pem=0.0, similarity=29.117863]
151: 5058739 Sweden
[pem_pos=0, final_scores=0.77256423,
final_local_scores=0.79854393, pem=-0.2135508,
similarity=29.081179]
152: 22577800 Stéphane_Simian
[pem_pos=0, final_scores=0.66542625,
final_local_scores=0.7996229, pem=0.0, similarity=27.952055]
153: 5843419 France
[pem_pos=0, final_scores=0.738103,
final_local_scores=0.75028306, pem=-0.16240834,
similarity=27.362858]
154: 3943969 Mikael_Tillström
[pem_pos=0, final_scores=0.643352,
final_local_scores=0.75433993, pem=0.0, similarity=26.600721]
155: 5058739 Sweden
[pem_pos=0, final_scores=0.7826008,
final_local_scores=0.8095074, pem=-0.2135508,
similarity=29.408352]
156: 8087628 Egypt
[pem_pos=0, final_scores=0.5888023,
final_local_scores=0.64557093, pem=-0.27480945,
similarity=24.849295]

157: 1723684 Roberto_Carretero
[pem_pos=0, final_scores=0.43284476,
final_local_scores=0.53371185, pem=0.0, similarity=20.016739]
158: 26667 Spain
[pem_pos=0, final_scores=0.62463427,
final_local_scores=0.63168013, pem=-0.18857051,
similarity=23.965784]
159: 14891740 Jordi_Burillo
[pem_pos=0, final_scores=0.5954705,
final_local_scores=0.71230024, pem=0.0, similarity=25.346172]
160: 26667 Spain
[pem_pos=0, final_scores=0.6377328,
final_local_scores=0.64598835, pem=-0.18857051,
similarity=24.392769]
161: 1408278 Thomas_Johansson
[pem_pos=0, final_scores=0.7714874,
final_local_scores=0.85558045, pem=0.0, similarity=29.62194]
162: 5058739 Sweden
[pem_pos=0, final_scores=0.80518323,
final_local_scores=0.8341752, pem=-0.2135508,
similarity=30.144487]
163: 3929800 Renzo_Furlan
[pem_pos=0, final_scores=0.6582495,
final_local_scores=0.7631305, pem=0.0, similarity=26.863049]
164: 14532 Italy
[pem_pos=0, final_scores=0.7207846,
final_local_scores=0.75892556, pem=-0.20612453,
similarity=27.858503]
165: 2261491 Mark_Knowles
[pem_pos=0, final_scores=0.5723149,
final_local_scores=0.661006, pem=0.0, similarity=23.815449]
166: 3451 The_Bahamas
[pem_pos=0, final_scores=0.48952574,
final_local_scores=0.52909654, pem=-0.35113433,
similarity=21.788532]
167: 2936118 Marcelo_Filippini
[pem_pos=0, final_scores=0.7650953,
final_local_scores=0.85880625, pem=0.0, similarity=29.718204]
168: 31843 Uruguay
[pem_pos=0, final_scores=0.43229407,
final_local_scores=0.428136, pem=-0.44594222,
similarity=19.291248]
169: 3056059 Jared_Palmer
[pem_pos=0, final_scores=0.70510024,
final_local_scores=0.82874686, pem=0.0, similarity=28.821173]
170: 3434750 United_States
[pem_pos=0, final_scores=0.51020473,
final_local_scores=0.50974804, pem=-0.14075686,
similarity=20.067068]
171: 1717291 Marc_Rosset
[pem_pos=0, final_scores=0.7013976,
final_local_scores=0.79987454, pem=0.0, similarity=27.959564]

172: 26748 Switzerland
[pem_pos=0, final_scores=0.65861624,
final_local_scores=0.66409355, pem=-0.18598008,
similarity=24.918978]
173: 2634301 Amy_Frazier
[pem_pos=0, final_scores=0.6159261,
final_local_scores=0.69639826, pem=0.0, similarity=24.871626]
174: 3434750 United_States
[pem_pos=0, final_scores=0.5125608,
final_local_scores=0.51232165, pem=-0.14075686,
similarity=20.143871]
175: 3877687 Larisa_Neiland
[pem_pos=0, final_scores=0.83857304,
final_local_scores=0.9600534, pem=0.0, similarity=32.73962]
176: 17514 Latvia
[pem_pos=0, final_scores=0.6316467,
final_local_scores=0.64723194, pem=-0.35144013,
similarity=25.315592]
177: 2199290 Lisa_Raymond
[pem_pos=0, final_scores=0.69923466,
final_local_scores=0.7835544, pem=0.0, similarity=27.472538]
178: 3434750 United_States
[pem_pos=0, final_scores=0.47296804,
final_local_scores=0.46907264, pem=-0.14075686,
similarity=18.853233]
179: 6426781 Lori_McNeil
[pem_pos=0, final_scores=0.6947852,
final_local_scores=0.778846, pem=0.0, similarity=27.332031]
180: 3434750 United_States
[pem_pos=0, final_scores=0.5025205,
final_local_scores=0.50135416, pem=-0.14075686,
similarity=19.81658]
181: 26964606 Austria
[pem_pos=0, final_scores=0.6589039,
final_local_scores=0.6637283, pem=-0.25262418,
similarity=25.270498]
182: 1028356 Zina_Garrison
[pem_pos=0, final_scores=0.54511464,
final_local_scores=0.6228145, pem=0.0, similarity=22.67574]
183: 3434750 United_States
[pem_pos=0, final_scores=0.5288705,
final_local_scores=0.5301376, pem=-0.14075686,
similarity=20.675533]
184: 784152 Conchita_Martínez
[pem_pos=0, final_scores=0.85572857,
final_local_scores=0.95740956, pem=0.0, similarity=32.660725]
185: 26667 Spain
[pem_pos=0, final_scores=0.6566585,
final_local_scores=0.66666174, pem=-0.18857051,
similarity=25.009705]
186: 4038121 Ruxandra_Dragomir
[pem_pos=0, final_scores=0.78685445,

```
final_local_scores=0.90054446, pem=0.0, similarity=30.963757]
187: 25445 Romania
    [pem_pos=0, final_scores=0.7580138,
final_local_scores=0.78160745, pem=-0.21241479,
similarity=28.569582]
188: 2631927 Naoko_Sawamatsu
    [pem_pos=0, final_scores=0.540727,
final_local_scores=0.6423971, pem=0.0, similarity=23.260124]
189: 15573 Japan
    [pem_pos=0, final_scores=0.7774163,
final_local_scores=0.83171064, pem=-0.13549899,
similarity=29.64648]
190: 2482750 Rennae_Stubbs
    [pem_pos=0, final_scores=0.63769335,
final_local_scores=0.7366106, pem=0.0, similarity=26.071642]
191: 4689264 Australia
    [pem_pos=0, final_scores=0.76530415,
final_local_scores=0.807039, pem=-0.17168477,
similarity=29.107016]
192: 2527510 Miriam_Oremans
    [pem_pos=0, final_scores=0.745626,
final_local_scores=0.8839149, pem=0.0, similarity=30.467495]
193: 21148 Netherlands
    [pem_pos=0, final_scores=0.61796474,
final_local_scores=0.6258494, pem=-0.24940795,
similarity=24.122627]
194: 15443594 Radka_Zrubáková
    [pem_pos=0, final_scores=0.53254384,
final_local_scores=0.6416746, pem=0.0, similarity=23.238562]
195: 26830 Slovakia
    [pem_pos=0, final_scores=0.66298825,
final_local_scores=0.6683778, pem=-0.36602598,
similarity=26.025948]
196: 4025956 Doug_Flach
    [pem_pos=0, final_scores=0.48798537,
final_local_scores=0.60211533, pem=0.0, similarity=22.058035]
197: 3434750 United_States
    [pem_pos=0, final_scores=0.4990551,
final_local_scores=0.4975688, pem=-0.14075686,
similarity=19.703615]
198: 4026677 Gianluca_Pozzi
    [pem_pos=0, final_scores=0.5811445,
final_local_scores=0.6961982, pem=0.0, similarity=24.865656]
199: 14532 Italy
    [pem_pos=0, final_scores=0.76107234,
final_local_scores=0.80293375, pem=-0.20612453,
similarity=29.171795]
200: 979076 Cédric_Pioline
    [pem_pos=0, final_scores=0.815067,
final_local_scores=0.9015146, pem=0.0, similarity=30.992708]
201: 5843419 France
    [pem_pos=0, final_scores=0.7281768,
```

```

final_local_scores=0.73944026, pem=-0.16240834,
similarity=27.039286]
202: 4076952 Francisco_Clavet
    [pem_pos=0, final_scores=0.645689,
final_local_scores=0.7556067, pem=0.0, similarity=26.638525]
203: 26667 Spain
    [pem_pos=0, final_scores=0.62459713,
final_local_scores=0.6316396, pem=-0.18857051,
similarity=23.964575]
204: 420527 Javier_Sánchez
    [pem_pos=0, final_scores=0.54288644,
final_local_scores=0.6284099, pem=0.0, similarity=22.842718]
205: 26667 Spain
    [pem_pos=0, final_scores=0.62055206,
final_local_scores=0.627221, pem=-0.18857051,
similarity=23.832714]
206: 12871566 David_Škoch
    [pem_pos=0, final_scores=0.76748616,
final_local_scores=0.8898713, pem=0.0, similarity=30.645247]
207: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.60764116,
final_local_scores=0.6090372, pem=-0.27794555,
similarity=23.77611]
208: 164640 Steffi_Graf
    [pem_pos=0, final_scores=0.8481346,
final_local_scores=0.9059817, pem=0.0, similarity=31.126015]
209: 11867 Germany
    [pem_pos=0, final_scores=0.8758319,
final_local_scores=0.90580636, pem=-0.15529114,
similarity=31.965279]
210: 2672548 Yayuk_Basuki
    [pem_pos=0, final_scores=0.77723163,
final_local_scores=0.88456887, pem=0.0, similarity=30.487013]
211: 14579 Indonesia
    [pem_pos=0, final_scores=0.5872737,
final_local_scores=0.6369366, pem=-0.23151328,
similarity=24.356178]

```

Sample: b'531_BASEBALL'

```

BASEBALL - [1 1 MAJOR LEAGUE 1] STANDINGS AFTER MONDAY 'S
GAMES . [2 2 NEW YORK 2] 1996-08-27 [3 3 Major League Baseball
3] standings after games played on Monday ( tabulate under
won , lost , winning percentage and games behind ) : [4 4 AME
RICAN LEAGUE EASTERN DIVISION 4] W L PCT GB [5 5 NEW YORK 5]
74 56 .569 - [6 6 BALTIMORE 6] 69 61 .531 5 [7 7 BOSTON 7] 67
65 .508 8 [8 8 TORONTO 8] 62 70 .470 13 [9 9 DETROIT 9] 47
84 .359 27 1/2 [10 10 CENTRAL DIVISION 10] [11 11 CLEVELAND 11
] 78 53 .595 - 1 CHICAGO 70 63 .526 9 [12 12 MINNESOTA 12] 65
66 .496 13 [13 13 MILWAUKEE 13] 63 69 .477 15 1/2 [14 14 KANS
AS CITY 14] 59 73 .447 19 1/2 [15 15 WESTERN DIVISION 15] [16
16 TEXAS 16] 75 56 .573 - [17 17 SEATTLE 17] 67 63 .515 7 1/2
[18 18 OAKLAND 18] 63 71 .470 13 1/2 2 CALIFORNIA 61 70 .466

```

14 TUESDAY , AUGUST 27TH SCHEDULE [19 19 CLEVELAND 19] AT [20
20 DETROIT 20] [21 21 OAKLAND 21] AT [22 22 BALTIMORE 22] [23
23 MINNESOTA 23] AT [1 3 TORONTO 1] [24 24 MILWAUKEE 24] AT 4
CHICAGO [25 25 TEXAS 25] AT [26 26 KANSAS CITY 26] [27 27 BOS
TON 27] AT 5 **CALIFORNIA** [2 6 NEW YORK 2] AT [28 28 SEATTLE 28]
[29 29 NATIONAL LEAGUE EASTERN DIVISION 29] W L PCT GB [30 30
ATLANTA 30] 81 48 .628 - [3 7 MONTREAL 3] 70 59 .543 11 8 **FLO**
RIDA 61 70 .466 21 [31 31 NEW YORK 31] 59 72 .450 23 [32 32 PH
ILADELPHIA 32] 53 79 .402 29 1/2 [33 33 CENTRAL DIVISION 33] [
34 34 HOUSTON 34] 70 62 .530 - ST LOUIS 69 62 .527 1/2 9 **CHIC**
AGO 64 64 .500 4 [35 35 CINCINNATI 35] 64 66 .492 5 [36 36 PI
TTSBURGH 36] 55 75 .423 14 [37 37 WESTERN DIVISION 37] [38 38
SAN DIEGO 38] 72 60 .545 - 10 **LOS ANGELES** 70 60 .538 1 [39 39
COLORADO 39] 69 63 .523 3 [40 40 SAN FRANCISCO 40] 56 73 .434
14 1/2 TUESDAY , AUGUST 27TH SCHEDULE [41 41 PHILADELPHIA 41]
AT [42 42 SAN FRANCISCO 42] 11 **LOS ANGELES** AT [43 43 MONTREAL
43] [44 44 ATLANTA 44] AT [45 45 PITTSBURGH 45] [46 46 SAN DI
EGO 46] AT [47 47 NEW YORK 47] 12 **CHICAGO** AT [48 48 HOUSTON 48
] 13 **FLORIDA** AT ST LOUIS [49 49 CINCINNATI 49] AT [50 50 COLO
RADO 50]

1: 38776 Major_League_Baseball
2: 645042 New_York_City
3: 38776 Major_League_Baseball
4: 2004582 American_League_East
5: 4848143 New_York_Yankees
6: 4340 Baltimore_Orioles
7: 4339 Boston_Red_Sox
8: 30847 Toronto_Blue_Jays
9: 8579 Detroit_Tigers
10: 2456377 National_League_Central
11: 6652 Cleveland_Indians
12: 20050 Minnesota_Twins
13: 20152 Milwaukee_Brewers
14: 16883 Kansas_City_Royals
15: 2456474 National_League_West
16: 30857 Texas_Rangers_(baseball)
17: 27660 Seattle_Mariners
18: 22523 Oakland_Athletics
19: 6652 Cleveland_Indians
20: 8579 Detroit_Tigers
21: 22523 Oakland_Athletics
22: 4340 Baltimore_Orioles
23: 20050 Minnesota_Twins
24: 20152 Milwaukee_Brewers
25: 30857 Texas_Rangers_(baseball)
26: 16883 Kansas_City_Royals
27: 4339 Boston_Red_Sox
28: 27660 Seattle_Mariners
29: 2456294 National_League_East
30: 2140 Atlanta_Braves
31: 4848143 New_York_Yankees
32: 23741 Philadelphia_Phillies

```
33: 2456377 National_League_Central
34: 13894 Houston_Astros
35: 6671 Cincinnati_Reds
36: 24270 Pittsburgh_Pirates
37: 2456474 National_League_West
38: 28417 San_Diego_Padres
39: 6669 Colorado_Rockies
40: 28416 San_Francisco_Giants
41: 23741 Philadelphia_Phillies
42: 28416 San_Francisco_Giants
43: 7954681 Montreal
44: 2140 Atlanta_Braves
45: 24270 Pittsburgh_Pirates
46: 28417 San_Diego_Padres
47: 4848143 New_York_Yankees
48: 13894 Houston_Astros
49: 6671 Cincinnati_Reds
50: 6669 Colorado_Rockies
1: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.7343204, final_local_scores=-0.8265983,
pem=-4.591283, similarity=4.390481|
     RECALL MISS: 5945 Chicago_White_Sox]
2: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 7590031 California_Golden_Bears_football,
pem_pos=1, final_scores=-0.9125795,
final_local_scores=-1.0025678, pem=-4.4771376,
similarity=-1.4815378|
     RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
3: 64646 Toronto
    [winner: 30847 Toronto_Blue_Jays, pem_pos=6,
final_scores=0.15933065, final_local_scores=0.06785762,
pem=-4.422637, similarity=30.165707|
     gt: 64646 Toronto, pem_pos=0, final_scores=0.074748024,
final_local_scores=0.10161638, pem=-0.12291032,
similarity=7.790554 ]
4: 5945 Chicago_White_Sox
    [winner: 6886 Chicago, pem_pos=0,
final_scores=-0.41825384, final_local_scores=-0.4618397,
pem=-0.20658566, similarity=-8.56906|
     RECALL MISS: 5945 Chicago_White_Sox]
5: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 5407 California, pem_pos=0,
final_scores=-0.758134, final_local_scores=-0.7912098,
pem=-0.08662026, similarity=-19.05051|
     RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
6: 645042 New_York_City
    [winner: 4848143 New_York_Yankees, pem_pos=11,
final_scores=0.18042888, final_local_scores=0.083753645,
pem=-5.0692368, similarity=34.156387|
     gt: 645042 New_York_City, pem_pos=1,
final_scores=0.06783883, final_local_scores=0.09571284,
```

```
pem=-1.7906687, similarity=16.68391 ]
7: 7954681 Montreal
    [winner: 20153 Montreal_Expos, pem_pos=8,
final_scores=0.6710517, final_local_scores=0.6243321,
pem=-4.5098605, similarity=47.24635]
    gt: 7954681 Montreal, pem_pos=0,
final_scores=-0.3744514, final_local_scores=-0.37293315,
pem=-0.11023275, similarity=-6.4398937 ]
8: 11278 Miami_Marlins
    [winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.73276967, final_local_scores=-0.80337113,
pem=-4.620281, similarity=5.241323|
     RECALL MISS: 11278 Miami_Marlins]
9: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.74672985, final_local_scores=-0.8401537,
pem=-4.591283, similarity=3.98596|
     RECALL MISS: 5945 Chicago_White_Sox]
10: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 18213 Los_Angeles_Dodgers, pem_pos=15,
final_scores=-0.15807731, final_local_scores=-0.287782,
pem=-4.7524934, similarity=21.346521|
     RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
11: 18110 Los_Angeles
    [winner: 18110 Los_Angeles, pem_pos=0,
final_scores=0.02891843, final_local_scores=0.045381024,
pem=-0.100652285, similarity=5.991336|
     gt: 18110 Los_Angeles, pem_pos=0,
final_scores=0.02891843, final_local_scores=0.045381024,
pem=-0.100652285, similarity=5.991336 ]
12: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.6329404, final_local_scores=-0.715856,
pem=-4.591283, similarity=7.6952515|
     RECALL MISS: 5945 Chicago_White_Sox]
13: 11278 Miami_Marlins
    [winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.6902805, final_local_scores=-0.7569582,
pem=-4.620281, similarity=6.6263776|
     RECALL MISS: 11278 Miami_Marlins]
1: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.594186,
final_local_scores=0.5467371, pem=-0.47792417,
similarity=23.004463]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.37867433,
final_local_scores=0.43522036, pem=-1.7906687,
similarity=26.81549]
3: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.5413267,
final_local_scores=0.48899627, pem=0.0, similarity=18.682337]
4: 2004582 American_League_East
```

```
[pem_pos=0, final_scores=0.475213,
final_local_scores=0.4746613, pem=0.0, similarity=18.254553]
5: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.3147952,
final_local_scores=0.23052841, pem=-5.0692368,
similarity=38.536438]
6: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.2921832,
final_local_scores=1.2967331, pem=-4.046561,
similarity=64.79264]
7: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.1713116,
final_local_scores=1.168019, pem=-4.418931,
similarity=62.976566]
8: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.46747905,
final_local_scores=0.40455937, pem=-4.422637,
similarity=40.21356]
9: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.2857739,
final_local_scores=1.2901187, pem=-4.1401463,
similarity=65.10419]
10: 2456377 National_League_Central
    [pem_pos=0, final_scores=0.5757876,
final_local_scores=0.5951262, pem=-0.98951703,
similarity=27.230614]
11: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.2775445,
final_local_scores=1.2798865, pem=-4.191953,
similarity=65.080574]
12: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=1.2509608,
final_local_scores=1.2542818, pem=-4.2906566,
similarity=64.85324]
13: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.5392981,
final_local_scores=0.47890574, pem=-3.40205,
similarity=36.88209]
14: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.8097151,
final_local_scores=0.7744616, pem=-3.5915844,
similarity=46.732788]
15: 2456474 National_League_West
    [pem_pos=1, final_scores=0.56030136,
final_local_scores=0.5656292, pem=-1.2952092,
similarity=28.012764]
16: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.2952574,
final_local_scores=1.3038898, pem=-4.928425,
similarity=69.801926]
17: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.7486593,
```

```
final_local_scores=0.71267277, pem=-3.8567436,
similarity=46.33086]
18: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.374577,
final_local_scores=1.3875852, pem=-3.9188845,
similarity=66.809525]
19: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.112711,
final_local_scores=1.0998306, pem=-4.191953,
similarity=59.707348]
20: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=0.9863393,
final_local_scores=0.9630318, pem=-4.1401463,
similarity=55.343266]
21: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.2869016,
final_local_scores=1.2918131, pem=-3.9188845,
similarity=63.951496]
22: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.1571736,
final_local_scores=1.1492558, pem=-4.046561,
similarity=60.39162]
23: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=1.2293725,
final_local_scores=1.2306998, pem=-4.2906566,
similarity=64.149506]
24: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.5344555,
final_local_scores=0.47361594, pem=-3.40205,
similarity=36.724236]
25: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.1440234,
final_local_scores=1.1386895, pem=-4.928425,
similarity=64.872025]
26: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.80844635,
final_local_scores=0.7730757, pem=-3.5915844,
similarity=46.69143]
27: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.1135368,
final_local_scores=1.104909, pem=-4.418931,
similarity=61.093235]
28: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.72649056,
final_local_scores=0.6884567, pem=-3.8567436,
similarity=45.608204]
29: 2456294 National_League_East
    [pem_pos=0, final_scores=0.4293537,
final_local_scores=0.42349297, pem=0.0, similarity=16.727589]
30: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=1.1565468,
final_local_scores=1.1481411, pem=-4.136474,
```

```
similarity=60.84732]
31: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.56366414,
final_local_scores=0.50238866, pem=-5.0692368,
similarity=46.649292]
32: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.268319,
final_local_scores=1.2720649, pem=-4.5189333,
similarity=66.62533]
33: 2456377 National_League_Central
    [pem_pos=0, final_scores=0.58044463,
final_local_scores=0.6002132, pem=-0.98951703,
similarity=27.382421]
34: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.3287194,
final_local_scores=1.3383062, pem=-4.650927,
similarity=69.31991]
35: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.1844071,
final_local_scores=1.1803339, pem=-3.771247,
similarity=59.821857]
36: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.3196533,
final_local_scores=1.325988, pem=-4.03702,
similarity=65.613785]
37: 2456474 National_League_West
    [pem_pos=1, final_scores=0.5872031,
final_local_scores=0.5950152, pem=-1.2952092,
similarity=28.889702]
38: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.9434388,
final_local_scores=0.92184764, pem=-3.9307637,
similarity=52.97559]
39: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.2198417,
final_local_scores=1.2268534, pem=-4.482383,
similarity=65.077354]
40: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.93945116,
final_local_scores=0.91114837, pem=-4.409061,
similarity=55.257355]
41: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.1923356,
final_local_scores=1.1890645, pem=-4.5189333,
similarity=64.14843]
42: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.923538,
final_local_scores=0.8937656, pem=-4.409061,
similarity=54.73862]
43: 7954681 Montreal
    [pem_pos=0, final_scores=0.34192127,
final_local_scores=0.41474283, pem=-0.11023275,
```

```
similarity=17.06593]
44: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=0.9854196,
final_local_scores=0.96121055, pem=-4.136474,
similarity=55.268944]
45: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.2019188,
final_local_scores=1.197381, pem=-4.03702,
similarity=61.775894]
46: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.9634116,
final_local_scores=0.94366485, pem=-3.9307637,
similarity=53.62666]
47: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.34138864,
final_local_scores=0.2595864, pem=-5.0692368,
similarity=39.403587]
48: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.4635849,
final_local_scores=1.4856262, pem=-4.650927,
similarity=73.71623]
49: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.1757761,
final_local_scores=1.1709058, pem=-3.771247,
similarity=59.540504]
50: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=0.94978493,
final_local_scores=0.9318573, pem=-4.482383,
similarity=56.274086]
1: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.15933065,
final_local_scores=0.06785762, pem=-4.422637,
similarity=30.165707]
2: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.18042888,
final_local_scores=0.083753645, pem=-5.0692368,
similarity=34.156387]
3: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.6710517,
final_local_scores=0.6243321, pem=-4.5098605,
similarity=47.24635]
```

Sample: b'532_BASEBALL'

BASEBALL - [1 1 GIANTS 1] EDGE [2 2 PHILLIES 2] 1-0 . [3 3 SA
N FRANCISCO 3] 1996-08-27 William VanLandingham pitched eight
scoreless innings and [4 4 Glenallen Hill 4] drove in the
game 's only run with a first-inning single as the [5 5 San Fr
ancisco Giants 5] claimed a 1-0 victory over the [6 6 Philade
lphia Phillies 6] on Monday . VanLandingham (8-13) , who
entered the game with one complete game in the first 56
starts of his career , limited the [7 7 Phillies 7] to two
hits and two walks with four strikeouts . " We 've been

working all year on my follow-through , and I really concentrated on that , " VanLandingham said . " It gave me more life in all of my pitches , so the ball moved more . " At [8 8 Colorado 8] , [9 9 Andres Galarraga 9] homered and drove in three runs as the [10 10 Colorado Rockies 10] had 10 extra-base hits and Billy Swift won his first game in almost a year in a 9-5 rain-shortened seven-inning victory over the [11 11 Cincinnati Reds 11] . Swift (1-0) , who made his first start since June 3rd and underwent arthroscopic surgery on his right shoulder earlier in the season , allowed five runs and six hits in five innings . In [12 12 Houston 12] , [13 13 Andy Benes 13] allowed two runs over seven innings and [14 14 Royce Clayton 14] had a run-scoring single in the seventh to lift the [15 15 St. Louis Cardinals 15] to a 3-2 victory over the [16 16 Houston Astros 16] . [17 17 Benes 17] (14-9) allowed five hits , walked five and struck out 10 for his 11th win in 12 decisions . The [18 18 Cardinals 18] moved within one-half game of first-place [19 19 Houston 19] in the [20 20 National League Central Division 20] .

```
1: 28416 San_Francisco_Giants
2: 23741 Philadelphia_Phillies
3: 49728 San_Francisco
4: 1823095 Glenallen_Hill
5: 28416 San_Francisco_Giants
6: 23741 Philadelphia_Phillies
7: 23741 Philadelphia_Phillies
8: 5399 Colorado
9: 762127 Andrés_Galarraga
10: 6669 Colorado_Rockies
11: 6671 Cincinnati_Reds
12: 13774 Houston
13: 4335606 Andy_Benes
14: 4551616 Royce_Clayton
15: 23408844 St._Louis_Cardinals
16: 13894 Houston_Astros
17: 4335606 Andy_Benes
18: 23408844 St._Louis_Cardinals
19: 13774 Houston
20: 2456377 National_League_Central
1: 28416 San_Francisco_Giants
    [pem_pos=0, final_scores=0.5461073,
final_local_scores=0.4863801, pem=-0.9105524,
similarity=23.55599]
2: 23741 Philadelphia_Phillies
    [pem_pos=0, final_scores=0.48225522,
final_local_scores=0.41818023, pem=-0.11780777,
similarity=17.209702]
3: 49728 San_Francisco
    [pem_pos=0, final_scores=0.34328705,
final_local_scores=0.3890046, pem=-0.09025385,
similarity=16.189201]
4: 1823095 Glenallen_Hill
```

```
[pem_pos=0, final_scores=0.790397,
final_local_scores=0.84478426, pem=0.0, similarity=29.29976]
5: 28416 San_Francisco_Giants
[pem_pos=0, final_scores=0.61913437,
final_local_scores=0.5661511, pem=-0.07570464,
similarity=21.396484]
6: 23741 Philadelphia_Phillies
[pem_pos=0, final_scores=0.44622475,
final_local_scores=0.37882245, pem=-0.09993234,
similarity=15.937978]
7: 23741 Philadelphia_Phillies
[pem_pos=0, final_scores=0.5192089,
final_local_scores=0.45854652, pem=-0.11780777,
similarity=18.414312]
8: 5399 Colorado
[pem_pos=0, final_scores=0.35579646,
final_local_scores=0.4426552, pem=-0.26926568,
similarity=18.763737]
9: 762127 Andrés_Galarraga
[pem_pos=0, final_scores=0.53733927,
final_local_scores=0.5603086, pem=0.0, similarity=20.810438]
10: 6669 Colorado_Rockies
[pem_pos=0, final_scores=0.4804691,
final_local_scores=0.42466086, pem=-0.10158596,
similarity=17.31488]
11: 6671 Cincinnati_Reds
[pem_pos=0, final_scores=0.5148827,
final_local_scores=0.45754755, pem=-0.106723964,
similarity=18.324224]
12: 13774 Houston
[pem_pos=0, final_scores=0.69259614,
final_local_scores=0.7509367, pem=-0.3232624,
similarity=28.257113]
13: 4335606 Andy_Benes
[pem_pos=0, final_scores=0.5885051,
final_local_scores=0.61003625, pem=0.0, similarity=22.29441]
14: 4551616 Royce_Clayton
[pem_pos=0, final_scores=0.70873094,
final_local_scores=0.7609269, pem=0.0, similarity=26.797289]
15: 23408844 St._Louis_Cardinals
[pem_pos=0, final_scores=0.44477266,
final_local_scores=0.3799165, pem=-0.2366796,
similarity=16.71428]
16: 13894 Houston_Astros
[pem_pos=0, final_scores=0.38978547,
final_local_scores=0.31836107, pem=-0.049621258,
similarity=13.86009]
17: 4335606 Andy_Benes
[pem_pos=0, final_scores=0.49418837,
final_local_scores=0.5176866, pem=0.0, similarity=19.538513]
18: 23408844 St._Louis_Cardinals
[pem_pos=0, final_scores=0.56749123,
```

```
final_local_scores=0.5139679, pem=-1.1265292,
similarity=25.55378]
19: 13774 Houston
    [pem_pos=0, final_scores=0.47455126,
final_local_scores=0.51275575, pem=-0.3232624,
similarity=21.149319]
20: 2456377 National_League_Central
    [pem_pos=0, final_scores=0.49813473,
final_local_scores=0.52293825, pem=0.0, similarity=19.695234]
```

Sample: b'533_BASEBALL'

BASEBALL - [1 1 ORIOLES 1] WIN , [2 2 YANKEES 2] LOSE . [3 3 BALTIMORE 3] 1996-08-27 [4 4 Cal Ripken 4] 's bases-loaded walk scored [5 5 Brady Anderson 5] with the winning run in the bottom of the 10th as the [6 6 Baltimore Orioles 6] regained control of the top spot in the wild-card race with a wild 12-11 victory over the [7 7 Oakland Athletics 7] . Trailing by a run entering the 11th , the [8 8 Orioles 8] rallied against [9 9 Oakland 9] reliever [10 10 Mark Acre 10] (0-2) with a walk and a triple by [11 11 Brady Anderson 11] to tie the game . Then [12 12 Oakland 12] manager [13 13 Art Howe 13] decided to intentionally walk [14 14 Rafael Palmeiro 14] and [15 15 Bobby Bonilla 15] to load the bases but [16 16 A cre 16] was nowhere near the plate to [17 17 Ripken 17] . The decisive pitch nearly hit [18 18 Ripken 18] and gave the [19 19 Orioles 19] a one-half game lead over the [20 20 Chicago White Sox 20] in the wild-card race . In [21 21 Seattle 21] , [22 22 Jay Buhner 22] 's eighth-inning single snapped a tie as the [23 23 Seattle Mariners 23] edged the [24 24 New York Yankees 24] 2-1 in the opener of a three-game series . [25 25 New York 25] starter [26 26 Jimmy Key 26] left the game in the first inning after [27 27 Seattle 27] shortstop [28 28 Alex Rodriguez 28] lined a shot off his left elbow . The [29 29 Yankees 29] have lost 12 of their last 19 games and their lead in the AL [30 30 East 30] over [31 31 Baltimore 31] fell to five games . At [32 32 California 32] , [33 33 Tim Wakefield 33] pitched a six-hitter for his third complete game of the season and [34 34 Mo Vaughn 34] and Troy O'Leary hit solo home runs in the second inning as the surging [35 35 Boston Red Sox 35] won their third straight 4-1 over the [36 36 California Angels 36] . [37 37 Boston 37] has won seven of eight and is 20-6 since August 2nd . The [38 38 Red Sox 38] are two games over .500 for the first time this season . In [39 39 Chicago 39] , [40 40 Cal Eldred 40] pitched 5-1/3 scoreless innings and [41 41 John Jaha 41] scored one run and doubled in another as the [42 42 Milwaukee Brewers 42] held off the slumping [43 43 Chicago White Sox 43] , 3-2 . Eldred (6-5) walked one and struck out three . [44 44 Angel Miranda 44] retired one batter and [45 45 Bob Wickman 45] retired the next four but loaded the bases in the eighth . In [46 46 Detroit 46] , [47 47 Jim Thome 47] 's solo homer in the ninth inning snapped a tie and [48 48 Charles Nagy 48] pitched a

three-hitter for his first win in over a month , leading the [49 49 Cleveland Indians 49] to their 11th straight victory over the [50 50 Detroit Tigers 50] , 2-1 . With the score tied 1-1 in the ninth , Thome hit a 2-2 pitch from starter [51 51 Felipe Lira 51] (6-11) over the left-field fence for his 29th homer . In [52 52 Toronto 52] , [53 53 Juan Guzman 53] allowed three runs over seven innings to make homers by [54 54 Joe Carter 54] and [55 55 Carlos Delgado 55] stand up as the surging [56 56 Toronto Blue Jays 56] held off the [57 57 Minnesota Twins 57] , 5-3 . [58 58 Toronto 58] returned home from a 10-game road trip and won for the eighth time in nine games as Guzman (11-8) allowed nine hits and struck out eight without a walk .

- 1: 4340 Baltimore_Orioles
- 2: 4848143 New_York_Yankees
- 3: 26997138 Baltimore
- 4: 153856 Cal_Ripken,_Jr.
- 5: 390266 Brady_Anderson
- 6: 4340 Baltimore_Orioles
- 7: 22523 Oakland_Athletics
- 8: 4340 Baltimore_Orioles
- 9: 22523 Oakland_Athletics
- 10: 18370398 Mark_Acre
- 11: 390266 Brady_Anderson
- 12: 22523 Oakland_Athletics
- 13: 3342240 Art_Howe
- 14: 230185 Rafael_Palmeiro
- 15: 905899 Bobby_Bonilla
- 16: 18370398 Mark_Acre
- 17: 153856 Cal_Ripken,_Jr.
- 18: 153856 Cal_Ripken,_Jr.
- 19: 4340 Baltimore_Orioles
- 20: 5945 Chicago_White_Sox
- 21: 11388236 Seattle
- 22: 385914 Jay_Buhner
- 23: 27660 Seattle_Mariners
- 24: 4848143 New_York_Yankees
- 25: 4848143 New_York_Yankees
- 26: 385773 Jimmy_Key
- 27: 27660 Seattle_Mariners
- 28: 212472 Alex_Rodriguez
- 29: 4848143 New_York_Yankees
- 30: 2004582 American_League_East
- 31: 4340 Baltimore_Orioles
- 32: 5407 California
- 33: 331205 Tim_Wakefield
- 34: 395171 Mo_Vaughn
- 35: 4339 Boston_Red_Sox
- 36: 1360083 Los_Angeles_Angels_ofAnaheim
- 37: 4339 Boston_Red_Sox
- 38: 4339 Boston_Red_Sox
- 39: 6886 Chicago

```
40: 3279577 Cal_Eldred
41: 2787747 John_Jaha
42: 20152 Milwaukee_Brewers
43: 5945 Chicago_White_Sox
44: 18939866 Ángel_Miranda
45: 1567114 Bob_Wickman
46: 8687 Detroit
47: 726320 Jim_Thome
48: 2406184 Charles_Nagy
49: 6652 Cleveland_Indians
50: 8579 Detroit_Tigers
51: 698835 Felipe_Lira
52: 64646 Toronto
53: 3139393 Juan_Guzmán_(baseball)
54: 301648 Joe_Carter
55: 300684 Carlos_Delgado
56: 30847 Toronto_Blue_Jays
57: 20050 Minnesota_Twins
58: 30847 Toronto_Blue_Jays
1: 4340 Baltimore_Orioles
    [pem_pos=0, final_scores=0.67085594,
final_local_scores=0.6221847, pem=-0.18171716,
similarity=23.645151]
2: 4848143 New_York_Yankees
    [pem_pos=0, final_scores=0.5493185,
final_local_scores=0.488643, pem=-0.4294726,
similarity=21.007334]
3: 26997138 Baltimore
    [pem_pos=0, final_scores=0.5267142,
final_local_scores=0.59077823, pem=-0.19949137,
similarity=22.804579]
4: 153856 Cal_Ripken,_Jr.
    [pem_pos=0, final_scores=0.36400425,
final_local_scores=0.34511837, pem=0.0, similarity=14.388734]
5: 390266 Brady_Anderson
    [pem_pos=0, final_scores=0.56783146,
final_local_scores=0.5881615, pem=0.0, similarity=21.641626]
6: 4340 Baltimore_Orioles
    [pem_pos=0, final_scores=0.4478752,
final_local_scores=0.3786121, pem=-0.21811225,
similarity=16.57438]
7: 22523 Oakland_Athletics
    [pem_pos=0, final_scores=0.538776,
final_local_scores=0.4802599, pem=-0.06842491,
similarity=18.793732]
8: 4340 Baltimore_Orioles
    [pem_pos=0, final_scores=0.54424924,
final_local_scores=0.48388612, pem=-0.18171716,
similarity=19.518045]
9: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=0.98647314,
final_local_scores=0.96930104, pem=-3.9188845,
```

```
similarity=54.32709]
10: 18370398 Mark_Acre
    [pem_pos=0, final_scores=0.62827176,
final_local_scores=0.76032066, pem=0.0, similarity=26.779198]
11: 390266 Brady_Anderson
    [pem_pos=0, final_scores=0.55699056,
final_local_scores=0.57631946, pem=0.0, similarity=21.288235]
12: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=0.92969674,
final_local_scores=0.90728134, pem=-3.9188845,
similarity=52.4763]
13: 3342240 Art_Howe
    [pem_pos=0, final_scores=0.5107978,
final_local_scores=0.5364627, pem=0.0, similarity=20.098831]
14: 230185 Rafael_Palmeiro
    [pem_pos=0, final_scores=0.56222796,
final_local_scores=0.5814525, pem=0.0, similarity=21.441416]
15: 905899 Bobby_Bonilla
    [pem_pos=0, final_scores=0.67837244,
final_local_scores=0.71149516, pem=0.0, similarity=25.322147]
16: 18370398 Mark_Acre
    [pem_pos=0, final_scores=1.0539334,
final_local_scores=1.2315843, pem=0.0, similarity=40.842644]
17: 153856 Cal_Ripken,_Jr.
    [pem_pos=0, final_scores=0.3454948,
final_local_scores=0.3248996, pem=0.0, similarity=13.785365]
18: 153856 Cal_Ripken,_Jr.
    [pem_pos=0, final_scores=0.3680992,
final_local_scores=0.3495915, pem=0.0, similarity=14.522221]
19: 4340 Baltimore_Orioles
    [pem_pos=0, final_scores=0.5167675,
final_local_scores=0.45386648, pem=-0.18171716,
similarity=18.6222]
20: 5945 Chicago_White_Sox
    [pem_pos=0, final_scores=0.47565323,
final_local_scores=0.40836477, pem=-0.07886935,
similarity=16.705036]
21: 11388236 Seattle
    [pem_pos=0, final_scores=0.53424805,
final_local_scores=0.60721165, pem=-0.1054736,
similarity=22.783703]
22: 385914 Jay_Buhner
    [pem_pos=0, final_scores=0.56435865,
final_local_scores=0.5859784, pem=0.0, similarity=21.576477]
23: 27660 Seattle_Mariners
    [pem_pos=0, final_scores=0.4391045,
final_local_scores=0.3762648, pem=-0.008242749,
similarity=15.36303]
24: 4848143 New_York_Yankees
    [pem_pos=0, final_scores=0.6112876,
final_local_scores=0.55633485, pem=-0.14222674,
similarity=21.465305]
```

```
25: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.48380256,
final_local_scores=0.41740292, pem=-5.0692368,
similarity=44.113148]
26: 385773 Jimmy_Key
    [pem_pos=0, final_scores=0.6566531,
final_local_scores=0.6841549, pem=0.0, similarity=24.50626]
27: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.593834,
final_local_scores=0.54528326, pem=-3.8567436,
similarity=41.335625]
28: 212472 Alex_Rodriguez
    [pem_pos=0, final_scores=0.51181364,
final_local_scores=0.4986061, pem=0.0, similarity=18.969112]
29: 4848143 New_York_Yankees
    [pem_pos=0, final_scores=0.47395867,
final_local_scores=0.40632385, pem=-0.4294726,
similarity=18.550766]
30: 2004582 American_League_East
    [pem_pos=14, final_scores=0.37122303,
final_local_scores=0.361376, pem=-4.0594826,
similarity=36.94999]
31: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.0532954,
final_local_scores=1.0399418, pem=-4.046561,
similarity=57.12948]
32: 5407 California
    [pem_pos=0, final_scores=0.595946,
final_local_scores=0.69987917, pem=-0.08662026,
similarity=25.446558]
33: 331205 Tim_Wakefield
    [pem_pos=0, final_scores=0.70731515,
final_local_scores=0.709679, pem=0.0, similarity=25.26795]
34: 395171 Mo_Vaughn
    [pem_pos=0, final_scores=0.8625588,
final_local_scores=0.89719564, pem=0.0, similarity=30.86382]
35: 4339 Boston_Red_Sox
    [pem_pos=0, final_scores=0.44840103,
final_local_scores=0.3801545, pem=-0.08447198,
similarity=15.893653]
36: 1360083 Los_Angeles_Angels_of_Anheim
    [pem_pos=0, final_scores=0.5189027,
final_local_scores=0.45982033, pem=-0.016987266,
similarity=17.904049]
37: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=0.87815136,
final_local_scores=0.84959143, pem=-4.418931,
similarity=53.47405]
38: 4339 Boston_Red_Sox
    [pem_pos=0, final_scores=0.50995594,
final_local_scores=0.4473939, pem=-0.07716359,
similarity=17.860466]
```

```
39: 6886 Chicago
    [pem_pos=0, final_scores=0.6045241,
final_local_scores=0.66746897, pem=-0.20658566,
similarity=25.131763]
40: 3279577 Cal_Eldred
    [pem_pos=0, final_scores=0.57172924,
final_local_scores=0.62607235, pem=0.0, similarity=22.77296]
41: 2787747 John_Jaha
    [pem_pos=0, final_scores=0.61386794,
final_local_scores=0.69433653, pem=0.0, similarity=24.8101]
42: 20152 Milwaukee_Brewers
    [pem_pos=0, final_scores=0.44032615,
final_local_scores=0.3763156, pem=-0.12176443,
similarity=15.9818945]
43: 5945 Chicago_White_Sox
    [pem_pos=0, final_scores=0.4911738,
final_local_scores=0.42531872, pem=-0.07886935,
similarity=17.210976]
44: 18939866 Ángel_Miranda
    [pem_pos=0, final_scores=0.59250385,
final_local_scores=0.67623407, pem=0.0, similarity=24.269886]
45: 1567114 Bob_Wickman
    [pem_pos=0, final_scores=0.65954906,
final_local_scores=0.6979682, pem=0.0, similarity=24.918476]
46: 8687 Detroit
    [pem_pos=0, final_scores=0.54652584,
final_local_scores=0.6012818, pem=-0.21700948,
similarity=23.213293]
47: 726320 Jim_Thome
    [pem_pos=0, final_scores=0.46923703,
final_local_scores=0.4550197, pem=0.0, similarity=17.66841]
48: 2406184 Charles_Nagy
    [pem_pos=0, final_scores=0.68881476,
final_local_scores=0.72376347, pem=0.0, similarity=25.68826]
49: 6652 Cleveland_Indians
    [pem_pos=0, final_scores=0.58639395,
final_local_scores=0.5295511, pem=-0.08813761,
similarity=20.37188]
50: 8579 Detroit_Tigers
    [pem_pos=0, final_scores=0.55955774,
final_local_scores=0.50416285, pem=-0.08904677,
similarity=19.619188]
51: 698835 Felipe_Lira
    [pem_pos=0, final_scores=0.37762743,
final_local_scores=0.49127835, pem=0.0, similarity=18.750439]
52: 64646 Toronto
    [pem_pos=0, final_scores=0.64504445,
final_local_scores=0.7336024, pem=-0.12291032,
similarity=26.650278]
53: 3139393 Juan_Guzmán_(baseball)
    [pem_pos=0, final_scores=0.34927356,
final_local_scores=0.36880583, pem=0.0, similarity=15.095614]
```

```
54: 301648 Joe_Carter
    [pem_pos=0, final_scores=0.5939195,
final_local_scores=0.602632, pem=0.0, similarity=22.073454]
55: 300684 Carlos_Delgado
    [pem_pos=0, final_scores=0.546927,
final_local_scores=0.55682105, pem=0.0, similarity=20.706364]
56: 30847 Toronto_Blue_Jays
    [pem_pos=0, final_scores=0.46655208,
final_local_scores=0.40785903, pem=-0.018220792,
similarity=16.360126]
57: 20050 Minnesota_Twins
    [pem_pos=0, final_scores=0.5137519,
final_local_scores=0.4559918, pem=-0.052076127,
similarity=17.980616]
58: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.7426195,
final_local_scores=0.7094208, pem=-4.422637,
similarity=49.311234]
```

Sample: b'534_BASEBALL'

BASEBALL - [1 1 MAJOR LEAGUE 1] RESULTS MONDAY . [2 2 NEW YO
RK 2] Results of [3 3 Major League Baseball 3] games played
on Monday (home team in CAPS) : [4 4 American League 4] [5 5
Cleveland 5] 2 [6 6 DETROIT 6] 1 [7 7 BALTIMORE 7] 12 [8 8 Oa
kland 8] 11 (10 innings) [9 9 TORONTO 9] 5 [10 10 Minnesota
10] 3 [11 11 Milwaukee 11] 3 1 CHICAGO 2 [12 12 Boston 12] 4 2
CALIFORNIA 1 [13 13 SEATTLE 13] 2 [14 14 New York 14] 1 [15 15
National League 15] [16 16 SAN FRANCISCO 16] 1 [17 17 Philade
lphia 17] 0 St Louis 3 [18 18 HOUSTON 18] 2 [19 19 COLORADO 19
] 9 [20 20 Cincinnati 20] 5

1: 38776 Major_League_Baseball
2: 645042 New_York_City
3: 38776 Major_League_Baseball
4: 64777 American_League
5: 6652 Cleveland_Indians
6: 8579 Detroit_Tigers
7: 4340 Baltimore_Orioles
8: 22523 Oakland_Athletics
9: 30847 Toronto_Blue_Jays
10: 20050 Minnesota_Twins
11: 20152 Milwaukee_Brewers
12: 4339 Boston_Red_Sox
13: 27660 Seattle_Mariners
14: 4848143 New_York_Yankees
15: 64290 National_League
16: 28416 San_Francisco_Giants
17: 23741 Philadelphia_Phillies
18: 13894 Houston_Astros
19: 6669 Colorado_Rockies
20: 6671 Cincinnati_Reds
1: 5945 Chicago_White_Sox
 [winner: 73126 Chicago_Blackhawks, pem_pos=8,

```
final_scores=-0.75863785, final_local_scores=-0.8487549,
pem=-4.591283, similarity=3.729284]
    RECALL MISS: 5945 Chicago_White_Sox]
2: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 5407 California, pem_pos=0,
final_scores=-0.5181664, final_local_scores=-0.5216005,
pem=-0.08662026, similarity=-11.004831]
    RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
1: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.5726176,
final_local_scores=0.51958126, pem=-0.47792417,
similarity=22.194078]
2: 645042 New_York_City
    [pem_pos=2, final_scores=0.27798718,
final_local_scores=0.34504533, pem=-1.7906687,
similarity=24.124489]
3: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.44452512,
final_local_scores=0.3796597, pem=0.0, similarity=15.419517]
4: 64777 American_League
    [pem_pos=0, final_scores=0.55463004,
final_local_scores=0.52226555, pem=0.0, similarity=19.67516]
5: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=0.99414617,
final_local_scores=0.9689252, pem=-4.191953,
similarity=55.800865]
6: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=0.8570363,
final_local_scores=0.81990194, pem=-4.1401463,
similarity=51.071983]
7: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.2013911,
final_local_scores=1.1957974, pem=-4.046561,
similarity=61.78052]
8: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.2266241,
final_local_scores=1.2247, pem=-3.9188845,
similarity=61.948708]
9: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.6758819,
final_local_scores=0.63496065, pem=-4.422637,
similarity=47.08919]
10: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=0.7862608,
final_local_scores=0.74446195, pem=-4.2906566,
similarity=49.639202]
11: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.49003965,
final_local_scores=0.42533296, pem=-3.40205,
similarity=35.283375]
12: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.1221513,
```

```
final_local_scores=1.1144605, pem=-4.418931,
similarity=61.378273]
13: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.66880924,
final_local_scores=0.6276353, pem=-3.8567436,
similarity=43.79317]
14: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.4946857,
final_local_scores=0.4284492, pem=-5.0692368,
similarity=44.44279]
15: 64290 National_League
    [pem_pos=0, final_scores=0.58812183,
final_local_scores=0.55383706, pem=-0.07431644,
similarity=21.021461]
16: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.75409657,
final_local_scores=0.70801944, pem=-4.409061,
similarity=49.195583]
17: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=0.8932,
final_local_scores=0.86081135, pem=-4.5189333,
similarity=54.352703]
18: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.0202332,
final_local_scores=1.0025694, pem=-4.650927,
similarity=59.300854]
19: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.0565124,
final_local_scores=1.0507936, pem=-4.482383,
similarity=59.823387]
20: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.0601845,
final_local_scores=1.042841, pem=-3.771247,
similarity=55.718796]
```

Sample: b'535_TENNIS'

TENNIS - [1 1 CHANG 1] , [2 2 WASHINGTON 2] ADVANCE , TWO WOMEN 'S SEEDS FALL . Larry Fine [3 3 NEW YORK 3] 1996-08-26 [4 4 Michael Chang 4] is playing in his 10th [5 5 U.S. Open 5] and enjoying his highest seeding ever , but the 24-year-old [6 6 American 6] had to overcome a case of the jitters Monday before winning his first-round match on opening day . [7 7 Chang 7] , seeded second behind defending champion [8 8 Pete Sampras 8] , took two hours 40 minutes to defeat 186th-ranked [9 9 Jaime Oncins 9] of [10 10 Brazil 10] 3-6 6-1 6-0 7-6 , 8-6 in the tiebreaker . " I was pretty tight the whole match , " conceded [11 11 Chang 11] , one of the hottest players on tour this summer with a 16-2 record on hardcourts that included two titles and a runner-up finish . " Everyone has moments when they get tight . Hopefully , this will have been my nerves for the whole tournament . " Joining [12 12 Chang 12] into the second round was [13 13 Wimbledon 13] runner-up [14

14 MaliVai Washington 14] , the 11th seed , who also needed four sets to get past talented [15 15 Moroccan 15] [16 16 Karim Alami 16] 6-4 2-6 7-6 (7-5) 6-1 . [17 17 Washington 17] 's win was not comfortable , either . The 27-year-old [18 18 American 18] hurried off the Stadium Court for treatment of an upset stomach after his two and a half hour struggle against [19 19 Alami 19] . " Towards the end of my match my stomach felt like week-old sushi , " said [20 20 Washington 20] . " Maybe it was a combination of the heat and something I ate . " [21 21 Chang 21] and [22 22 Washington 22] were the only men 's seeds in action on a day that saw two seeded women 's players fall . [23 23 Australian Open 23] runner-up [24 24 Anke Huber 24] of [25 25 Germany 25] , the sixth seed , was undone by an unlucky draw that put her against 17th ranked [26 26 South African 26] [27 27 Amanda Coetzer 27] in her opening match . [28 28 Coetzer 28] claimed revenge for the semifinal defeat she suffered to [29 29 Huber 29] in [30 30 Melbourne 30] by taking a 6-1 2-6 6-2 victory . Last year 's [31 31 Wimledon 31] junior champion , [32 32 Aleksandra Olsza 32] of [33 33 Poland 33] , removed another seed from the draw by eliminating number 12 [34 34 Magdalena Maleeva 34] of [35 35 Bulgaria 35] 6-4 6-2 . Other men 's winners included a pair of former [36 36 Grand Slam 36] tournament champions whose victories set up a showdown in the second round . [37 37 Germany 37] 's [38 38 Michael Stich 38] , the 1991 [39 39 Wimbledon 39] champion , and two-time [40 40 French Open 40] winner [41 41 Sergi Bruguera 41] of [42 42 Spain 42] will face each other next after beating [43 43 German 43] [44 44 Tommy Haas 44] 6-3 1-6 6-1 7-5 , and [45 45 Belgian 45] Kris Goossens 6-2 6-0 7-6 (7-1) , respectively . Alex O'Brien , who scored his first professional title eight days ago in [46 46 New Haven 46] , advanced to the second round with a 6-4 1-6 6-4 6-3 win over [47 47 Ecuador 47] 's [48 48 Nicolas Lapentti 48] . [49 49 Wimbledon 49] bad boy [50 50 Jeff Tarango 50] caught a break Monday when he advanced after the retirement of [51 51 German 51] [52 52 Alex Radulescu 52] due to heat exhaustion . [53 53 Tarango 53] was leading 6-7 (5-7) 6-4 6-1 3-1 . [54 54 Chang 54] blamed breezy conditions for some of the erratic play in his match with [55 55 Oncins 55] , who had beaten him in the round of 32 at the 1992 [56 56 Barcelona 56] [57 57 Olympics 57] . [58 58 Chang 58] committed an untidy 53 unforced errors , though he made seven fewer than the [59 59 Brazilian 59] , who also walloped a woeful 24 double faults . The most deflating double fault came when [60 60 Oncins 60] was serving to force a fifth set , leading 6-4 in the tiebreaker . The set point came after confusion at the net on the point previous , which was awarded to [61 61 Oncins 61] after the two exchanged shots at close quarters at the net . First [62 62 Chang 62] approached the umpire , then [63 63 Oncins 63] , then [64 64 Chang 64] again before play finally resumed . " Those little seconds were like an hour to me , " said [65 65 Oncins 65] , who promptly fired two serves barely

contained by the baseline as he frittered away his best chance . [66 66 Chang 66] ran off the next three points to close out the match but for [67 67 Oncins 67] , the contest was a personal victory . The 26-year-old [68 68 Brazilian 68] had risen into the top 30 in 1992 . The next year a close friend was struck by a stray bullet while riding home in a car from a soccer game in [69 69 Sao Paulo 69] . He died slumped against [70 70 Oncins 70] , who subsequently lost interest in tennis . " Two months ago I started talking about quitting , " said [71 71 Oncins 71] , who decided to give it one last try and made it through the Open qualifying tournament last weekend . " I believe in my game again , " he said . At the end of the day , a spate of withdrawals from the tournament were announced . Eighth seed [72 72 Jim Courier 72] withdrew because of a bruised right knee , and 1988 Open champion [73 73 Mats Wilander 73] bowed out due to a groin pull , organisers said . Women 's ninth seed [74 74 Mary Joe Fernandez 74] pulled out because of tendinitis in her right wrist .

1: 391834 Michael_Chang
2: 108956 Washington,_D.C.
3: 645042 New_York_City
4: 391834 Michael_Chang
5: 82729 US_Open_(tennis)
6: 3434750 United_States
7: 391834 Michael_Chang
8: 65961 Pete_Sampras
9: 3931197 Jaime_Oncins
10: 3383 Brazil
11: 391834 Michael_Chang
12: 391834 Michael_Chang
13: 91195 The_Championships,_Wimbledon
14: 873616 MaliVai_Washington
15: 19291 Morocco
16: 3930942 Karim_Alami
17: 873616 MaliVai_Washington
18: 3434750 United_States
19: 3930942 Karim_Alami
20: 873616 MaliVai_Washington
21: 391834 Michael_Chang
22: 873616 MaliVai_Washington
23: 97851 Australian_Open
24: 1443478 Anke_Huber
25: 11867 Germany
26: 17416221 South_Africa
27: 1951841 Amanda_Coetzer
28: 1951841 Amanda_Coetzer
29: 1443478 Anke_Huber
30: 17306237 Melbourne
31: 91195 The_Championships,_Wimbledon
32: 22207361 Aleksandra_Olsza
33: 22936 Poland

34: 1111175 Magdalena_Maleeva
35: 3415 Bulgaria
36: 197638 Grand_Slam_(tennis)
37: 11867 Germany
38: 979088 Michael_Stich
39: 91195 The_Championships,_Wimbledon
40: 147724 French_Open
41: 913625 Sergi_Bruguera
42: 26667 Spain
43: 11867 Germany
44: 1118005 Tommy_Haas
45: 3343 Belgium
46: 53825 New_Haven,_Connecticut
47: 9334 Ecuador
48: 2817661 Nicolás_Lapentti
49: 91195 The_Championships,_Wimbledon
50: 1862751 Jeff_Tarango
51: 11867 Germany
52: 4026041 Alex_Rădulescu
53: 1862751 Jeff_Tarango
54: 391834 Michael_Chang
55: 3931197 Jaime_Oncins
56: 4443 Barcelona
57: 27566 Summer_Olympic_Games
58: 391834 Michael_Chang
59: 3383 Brazil
60: 3931197 Jaime_Oncins
61: 3931197 Jaime_Oncins
62: 391834 Michael_Chang
63: 3931197 Jaime_Oncins
64: 391834 Michael_Chang
65: 3931197 Jaime_Oncins
66: 391834 Michael_Chang
67: 3931197 Jaime_Oncins
68: 3383 Brazil
69: 390875 São_Paulo
70: 3931197 Jaime_Oncins
71: 3931197 Jaime_Oncins
72: 685780 Jim_Courier
73: 200446 Mats_Wilander
74: 978947 Mary_Joe_Fernández
1: 391834 Michael_Chang
 [pem_pos=0, final_scores=0.9821808,
final_local_scores=0.9873608, pem=0.0, similarity=33.554527]
2: 108956 Washington,_D.C.
 [pem_pos=1, final_scores=0.2858165,
final_local_scores=0.36440635, pem=-1.7496481,
similarity=24.479183]
3: 645042 New_York_City
 [pem_pos=1, final_scores=0.38220507,
final_local_scores=0.452765, pem=-1.7906687,
similarity=27.339058]

```
4: 391834 Michael_Chang
    [pem_pos=0, final_scores=0.86788505,
final_local_scores=0.86125267, pem=0.0, similarity=29.79121]
5: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.51100427,
final_local_scores=0.4972741, pem=-1.2370842,
similarity=25.65682]
6: 3434750 United_States
    [pem_pos=0, final_scores=0.48733115,
final_local_scores=0.55934757, pem=-0.12994486,
similarity=21.48842]
7: 391834 Michael_Chang
    [pem_pos=0, final_scores=1.3612473,
final_local_scores=1.4014335, pem=0.0, similarity=45.91128]
8: 65961 Pete_Sampras
    [pem_pos=0, final_scores=0.7657631,
final_local_scores=0.75865746, pem=0.0, similarity=26.729565]
9: 3931197 Jaime_Oncins
    [pem_pos=0, final_scores=0.703114,
final_local_scores=0.7417876, pem=0.0, similarity=26.226135]
10: 3383 Brazil
    [pem_pos=0, final_scores=0.7709647,
final_local_scores=0.8645449, pem=-0.21313342,
similarity=31.048512]
11: 391834 Michael_Chang
    [pem_pos=0, final_scores=1.3348037,
final_local_scores=1.3725479, pem=0.0, similarity=45.049282]
12: 391834 Michael_Chang
    [pem_pos=0, final_scores=1.3314409,
final_local_scores=1.3688745, pem=0.0, similarity=44.939663]
13: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.5665967,
final_local_scores=0.5476495, pem=-0.8505555,
similarity=25.058119]
14: 873616 MaliVai_Washington
    [pem_pos=0, final_scores=0.90643376,
final_local_scores=0.9517904, pem=0.0, similarity=32.49304]
15: 19291 Morocco
    [pem_pos=0, final_scores=0.69407505,
final_local_scores=0.8130631, pem=-0.2674848,
similarity=29.807762]
16: 3930942 Karim_Alami
    [pem_pos=0, final_scores=0.65241605,
final_local_scores=0.7315168, pem=0.0, similarity=25.919632]
17: 873616 MaliVai_Washington
    [pem_pos=0, final_scores=0.5997847,
final_local_scores=0.6168228, pem=0.0, similarity=22.496935]
18: 3434750 United_States
    [pem_pos=0, final_scores=0.4445653,
final_local_scores=0.51263237, pem=-0.12994486,
similarity=20.094347]
19: 3930942 Karim_Alami
```

```
[pem_pos=0, final_scores=0.5526028,
final_local_scores=0.622486, pem=0.0, similarity=22.665937]
20: 873616 MaliVai_Washington
    [pem_pos=0, final_scores=0.7102173,
final_local_scores=0.7374537, pem=0.0, similarity=26.096802]
21: 391834 Michael_Chang
    [pem_pos=0, final_scores=1.1747745,
final_local_scores=1.1977403, pem=0.0, similarity=39.832676]
22: 873616 MaliVai_Washington
    [pem_pos=0, final_scores=0.7406078,
final_local_scores=0.7706507, pem=0.0, similarity=27.087467]
23: 97851 Australian_Open
    [pem_pos=0, final_scores=0.4540277,
final_local_scores=0.43534052, pem=-0.3261775,
similarity=18.854946]
24: 1443478 Anke_Huber
    [pem_pos=0, final_scores=0.95483476,
final_local_scores=1.0282195, pem=0.0, similarity=34.773834]
25: 11867 Germany
    [pem_pos=0, final_scores=0.822705,
final_local_scores=0.92069167, pem=-0.15529114,
similarity=32.409485]
26: 17416221 South_Africa
    [pem_pos=0, final_scores=0.54999,
final_local_scores=0.6680529, pem=-0.1389451,
similarity=24.78135]
27: 1951841 Amanda_Coetzer
    [pem_pos=0, final_scores=0.74911743,
final_local_scores=0.7970199, pem=0.0, similarity=27.874376]
28: 1951841 Amanda_Coetzer
    [pem_pos=0, final_scores=0.4996717,
final_local_scores=0.5245383, pem=0.0, similarity=19.742981]
29: 1443478 Anke_Huber
    [pem_pos=0, final_scores=0.8063994,
final_local_scores=0.8691762, pem=0.0, similarity=30.027664]
30: 17306237 Melbourne
    [pem_pos=0, final_scores=0.41325676,
final_local_scores=0.47709864, pem=-0.16841523,
similarity=19.243156]
31: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.56062496,
final_local_scores=0.5411263, pem=-0.8505555,
similarity=24.863455]
32: 22207361 Aleksandra_Olsza
    [pem_pos=0, final_scores=0.79548484,
final_local_scores=0.9452613, pem=0.0, similarity=32.298195]
33: 22936 Poland
    [pem_pos=0, final_scores=0.5899957,
final_local_scores=0.6668257, pem=-0.15333635,
similarity=24.822989]
34: 1111175 Magdalena_Maleeva
    [pem_pos=0, final_scores=0.8768286,
```

```
final_local_scores=0.93765706, pem=0.0, similarity=32.07127]
35: 3415 Bulgaria
    [pem_pos=0, final_scores=0.76713544,
final_local_scores=0.88855416, pem=-0.31727514,
similarity=32.331333]
36: 197638 Grand_Slam_(tennis)
    [pem_pos=0, final_scores=0.61448306,
final_local_scores=0.60238236, pem=-0.37449613,
similarity=24.102573]
37: 11867 Germany
    [pem_pos=0, final_scores=0.8232116,
final_local_scores=0.92124516, pem=-0.15529114,
similarity=32.426003]
38: 979088 Michael_Stich
    [pem_pos=0, final_scores=0.80070335,
final_local_scores=0.8417128, pem=0.0, similarity=29.208101]
39: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.48877245,
final_local_scores=0.46263832, pem=-0.8505555,
similarity=22.521215]
40: 147724 French_Open
    [pem_pos=0, final_scores=0.47830033,
final_local_scores=0.45847476, pem=-0.23693618,
similarity=19.060009]
41: 913625 Sergi_Bruguera
    [pem_pos=0, final_scores=0.8100365,
final_local_scores=0.8391656, pem=0.0, similarity=29.13209]
42: 26667 Spain
    [pem_pos=0, final_scores=0.5904071,
final_local_scores=0.6690227, pem=-0.18857051,
similarity=25.08016]
43: 11867 Germany
    [pem_pos=1, final_scores=0.7653937,
final_local_scores=0.85808784, pem=-1.0465823,
similarity=35.388245]
44: 1118005 Tommy_Haas
    [pem_pos=0, final_scores=0.871074,
final_local_scores=0.8936016, pem=0.0, similarity=30.756567]
45: 3343 Belgium
    [pem_pos=0, final_scores=0.7311907,
final_local_scores=0.80218816, pem=-0.12403572,
similarity=28.703133]
46: 53825 New_Haven,_Connecticut
    [pem_pos=0, final_scores=0.41682452,
final_local_scores=0.50361055, pem=-0.46266162,
similarity=21.634483]
47: 9334 Ecuador
    [pem_pos=0, final_scores=0.5474251,
final_local_scores=0.6710297, pem=-0.21404564,
similarity=25.278591]
48: 2817661 Nicolás_Lapentti
    [pem_pos=0, final_scores=0.7124269,
```

```
final_local_scores=0.72848123, pem=0.0, similarity=25.829046]
49: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.5039996,
final_local_scores=0.4792717, pem=-0.8505555,
similarity=23.01759]
50: 1862751 Jeff_Tarango
    [pem_pos=0, final_scores=0.6821154,
final_local_scores=0.72024417, pem=0.0, similarity=25.583235]
51: 11867 Germany
    [pem_pos=1, final_scores=0.7965537,
final_local_scores=0.8921254, pem=-1.0465823,
similarity=36.40399]
52: 4026041 Alex_Rădulescu
    [pem_pos=0, final_scores=0.67365927,
final_local_scores=0.7944328, pem=0.0, similarity=27.797173]
53: 1862751 Jeff_Tarango
    [pem_pos=0, final_scores=0.66552633,
final_local_scores=0.70212317, pem=0.0, similarity=25.042469]
54: 391834 Michael_Chang
    [pem_pos=0, final_scores=1.3179278,
final_local_scores=1.3541136, pem=0.0, similarity=44.499165]
55: 3931197 Jaime_Oncins
    [pem_pos=0, final_scores=0.55328923,
final_local_scores=0.5781268, pem=0.0, similarity=21.342169]
56: 4443 Barcelona
    [pem_pos=0, final_scores=0.4224612,
final_local_scores=0.45785606, pem=-0.37077636,
similarity=19.76939]
57: 27566 Summer_Olympic_Games
    [pem_pos=2, final_scores=0.25336796,
final_local_scores=0.2701193, pem=-2.7333555,
similarity=27.015024]
58: 391834 Michael_Chang
    [pem_pos=0, final_scores=1.2034734,
final_local_scores=1.2290895, pem=0.0, similarity=40.768196]
59: 3383 Brazil
    [pem_pos=0, final_scores=0.5720938,
final_local_scores=0.64730865, pem=-0.2733083,
similarity=24.892986]
60: 3931197 Jaime_Oncins
    [pem_pos=0, final_scores=0.5950236,
final_local_scores=0.62371534, pem=0.0, similarity=22.702623]
61: 3931197 Jaime_Oncins
    [pem_pos=0, final_scores=0.64454764,
final_local_scores=0.67781276, pem=0.0, similarity=24.316998]
62: 391834 Michael_Chang
    [pem_pos=0, final_scores=1.4186623,
final_local_scores=1.4641507, pem=0.0, similarity=47.782887]
63: 3931197 Jaime_Oncins
    [pem_pos=0, final_scores=0.68003374,
final_local_scores=0.716576, pem=0.0, similarity=25.47377]
64: 391834 Michael_Chang
```

```
[pem_pos=0, final_scores=1.254962,
final_local_scores=1.2853329, pem=0.0, similarity=42.44661]
65: 3931197 Jaime_Oncins
    [pem_pos=0, final_scores=0.718025,
final_local_scores=0.7580757, pem=0.0, similarity=26.712204]
66: 391834 Michael_Chang
    [pem_pos=0, final_scores=1.2366941,
final_local_scores=1.265378, pem=0.0, similarity=41.851116]
67: 3931197 Jaime_Oncins
    [pem_pos=0, final_scores=0.655467,
final_local_scores=0.6897405, pem=0.0, similarity=24.672945]
68: 3383 Brazil
    [pem_pos=0, final_scores=0.5861002,
final_local_scores=0.66260856, pem=-0.2733083,
similarity=25.349567]
69: 390875 São_Paulo
    [pem_pos=0, final_scores=0.33512038,
final_local_scores=0.37279117, pem=0.0, similarity=15.214545]
70: 3931197 Jaime_Oncins
    [pem_pos=0, final_scores=0.59758836,
final_local_scores=0.6265169, pem=0.0, similarity=22.786226]
71: 3931197 Jaime_Oncins
    [pem_pos=0, final_scores=0.79305595,
final_local_scores=0.8400356, pem=0.0, similarity=29.15805]
72: 685780 Jim_Courier
    [pem_pos=0, final_scores=0.78740007,
final_local_scores=0.79126644, pem=0.0, similarity=27.702682]
73: 200446 Mats_Wilander
    [pem_pos=0, final_scores=0.91133183,
final_local_scores=0.94110495, pem=0.0, similarity=32.17416]
74: 978947 Mary_Joe_Fernández
    [pem_pos=0, final_scores=0.5656051,
final_local_scores=0.60025257, pem=0.0, similarity=22.002447]
```

Sample: b'536_SOCCER'

```
SOCCER - [1 1 FRENCH 1] FIRST DIVISION SUMMARY . [2 2 PARIS 2
] 1996-08-27 Summary of a [3 3 French 3] first division
soccer match on Tuesday : [4 4 Auxerre 4] 0 [5 5 Marseille 5]
0 . Attendance : 20,000
1: 5843419 France
2: 22989 Paris
3: 5843419 France
4: 990753 AJ_Auxerre
5: 492068 Olympique_de_Marseille
1: 5843419 France
    [pem_pos=0, final_scores=0.71586865,
final_local_scores=0.70801747, pem=-0.80223274,
similarity=29.581034]
2: 22989 Paris
    [pem_pos=0, final_scores=0.57936066,
final_local_scores=0.5828245, pem=-0.115301244,
similarity=22.109385]
```

```
3: 5843419 France
    [pem_pos=0, final_scores=0.78967184,
final_local_scores=0.78863627, pem=-0.80223274,
similarity=31.986858]
4: 990753 AJ_Auxerre
    [pem_pos=0, final_scores=0.5598522,
final_local_scores=0.6033576, pem=-0.30518255,
similarity=23.754738]
5: 492068 Olympique_de_Marseille
    [pem_pos=1, final_scores=0.5532175,
final_local_scores=0.57143605, pem=-1.4345005,
similarity=28.94354]
```

Sample: b'537_SOCCER'

SOCCER - [1 1 MARSEILLE 1] HOLD [2 2 AUXERRE 2] TO GOALLESS DRAW . [3 3 PARIS 3] 1996-08-27 Former [4 4 European 4] champions [5 5 Marseille 5] held [6 6 French 6] champions [7 7 Auxerre 7] to a goalless draw in a lacklustre league match on Tuesday . The bill looked promising but both sides , struggling to find their form early in the season , were disappointing . [8 8 Auxerre 8] , who start their [9 9 Europe an Cup 9] campaign next week against [10 10 Ajax Amsterdam 10] , dominated the match but were unable to score . Unbeaten in four matches , they still trail leaders [11 11 Lens 11] by one point . [12 12 Lens 12] , who have won all their three league matches so far , host [13 13 Montpellier 13] on Wednesday night . Despite another dismal performance , especially in defence , [14 14 Marseille 14] restored some pride by keeping the reigning champions at bay after losing 2-1 at home to [15 15 Metz 15] last Saturday . After two seasons in the second division and after taking on half a dozen new recruits this season , some of whom do not speak a word of [16 16 French 16] , [17 17 Marseille 17] are not playing with any fluidity . But [18 18 German 18] international goalkeeper [19 19 Andreas Koepke 19] again proved a sound investment when under pressure from the [20 20 Auxerre 20] strikers , saving his team with a number of fine parries . [21 21 Marseille 21] now lie seventh in the league on five points .

```
1: 492068 Olympique_de_Marseille
2: 990753 AJ_Auxerre
3: 22989 Paris
4: 9239 Europe
5: 492068 Olympique_de_Marseille
6: 5843419 France
7: 990753 AJ_Auxerre
8: 990753 AJ_Auxerre
9: 44220 UEFA_Champions_League
10: 2273 AFC_Ajax
11: 1075697 RC_Lens
12: 1075697 RC_Lens
13: 1640865 Montpellier_HSC
```

```
14: 492068 Olympique_de_Marseille
15: 1135310 FC_Metz
16: 5843419 France
17: 492068 Olympique_de_Marseille
18: 11867 Germany
19: 2255217 Andreas_Kopke
20: 990753 AJ_Auxerre
21: 492068 Olympique_de_Marseille
1: 492068 Olympique_de_Marseille
    [pem_pos=1, final_scores=0.6132466,
final_local_scores=0.56097054, pem=-1.4345005,
similarity=28.63123]
2: 990753 AJ_Auxerre
    [pem_pos=0, final_scores=0.79077405,
final_local_scores=0.7692349, pem=-0.30518255,
similarity=28.704847]
3: 22989 Paris
    [pem_pos=0, final_scores=0.46884453,
final_local_scores=0.50815856, pem=-0.115301244,
similarity=19.881205]
4: 9239 Europe
    [pem_pos=0, final_scores=0.38365096,
final_local_scores=0.48705953, pem=-0.25557294,
similarity=20.014385]
5: 492068 Olympique_de_Marseille
    [pem_pos=1, final_scores=0.60488695,
final_local_scores=0.5518389, pem=-1.4345005,
similarity=28.358723]
6: 5843419 France
    [pem_pos=0, final_scores=0.5205615,
final_local_scores=0.5532246, pem=-0.80223274,
similarity=24.961706]
7: 990753 AJ_Auxerre
    [pem_pos=0, final_scores=0.58331203,
final_local_scores=0.54261416, pem=-0.30518255,
similarity=21.942032]
8: 990753 AJ_Auxerre
    [pem_pos=0, final_scores=0.48649937,
final_local_scores=0.43686098, pem=-0.30518255,
similarity=18.786146]
9: 44220 UEFA_Champions_League
    [pem_pos=0, final_scores=0.41316086,
final_local_scores=0.39494222, pem=-0.5339415,
similarity=18.779232]
10: 2273 AFC_Ajax
    [pem_pos=0, final_scores=0.6645989,
final_local_scores=0.65628135, pem=0.0, similarity=23.674458]
11: 1075697 RC_Lens
    [pem_pos=0, final_scores=0.510883,
final_local_scores=0.46933097, pem=-0.86081105,
similarity=22.77671]
12: 1075697 RC_Lens
```

```
[pem_pos=0, final_scores=0.4773063,
final_local_scores=0.4326536, pem=-0.86081105,
similarity=21.682184]
13: 1640865 Montpellier_HSC
    [pem_pos=1, final_scores=0.5667853,
final_local_scores=0.5348283, pem=-1.476352,
similarity=28.078688]
14: 492068 Olympique_de_Marseille
    [pem_pos=1, final_scores=0.6352784,
final_local_scores=0.5850369, pem=-1.4345005,
similarity=29.349417]
15: 1135310 FC_Metz
    [pem_pos=1, final_scores=0.6529967,
final_local_scores=0.6268825, pem=-1.9220247,
similarity=33.249405]
16: 5843419 France
    [pem_pos=0, final_scores=0.44456458,
final_local_scores=0.47020954, pem=-0.80223274,
similarity=22.48437]
17: 492068 Olympique_de_Marseille
    [pem_pos=1, final_scores=0.4753756,
final_local_scores=0.41036743, pem=-1.4345005,
similarity=24.136932]
18: 11867 Germany
    [pem_pos=1, final_scores=0.7849596,
final_local_scores=0.8838356, pem=-1.0465823,
similarity=36.15661]
19: 2255217 Andreas_Köpke
    [pem_pos=0, final_scores=0.9030177,
final_local_scores=1.0017278, pem=0.0, similarity=33.98327]
20: 990753 AJ_Auxerre
    [pem_pos=0, final_scores=0.4684983,
final_local_scores=0.41719753, pem=-0.30518255,
similarity=18.19935]
21: 492068 Olympique_de_Marseille
    [pem_pos=1, final_scores=0.71666604,
final_local_scores=0.67394054, pem=-1.4345005,
similarity=32.00248]
```

Sample: b'538_SOCCER'

```
SOCCER - [1 1 FRENCH 1] FIRST DIVISION RESULT . [2 2 PARIS 2]
1996-08-27 Result of a [3 3 French 3] first division soccer
match played on Tuesday : [4 4 Auxerre 4] 0 [5 5 Marseille 5]
0
1: 5843419 France
2: 22989 Paris
3: 5843419 France
4: 990753 AJ_Auxerre
5: 492068 Olympique_de_Marseille
1: 5843419 France
    [pem_pos=0, final_scores=0.7422939,
final_local_scores=0.736883, pem=-0.80223274,
```

```
similarity=30.442438]
2: 22989 Paris
    [pem_pos=0, final_scores=0.5751689,
final_local_scores=0.5782456, pem=-0.115301244,
similarity=21.97274]
3: 5843419 France
    [pem_pos=0, final_scores=0.79651123,
final_local_scores=0.79610723, pem=-0.80223274,
similarity=32.20981]
4: 990753 AJ_Auxerre
    [pem_pos=0, final_scores=0.572918,
final_local_scores=0.61763, pem=-0.30518255,
similarity=24.180655]
5: 492068 Olympique_de_Marseille
    [pem_pos=1, final_scores=0.47680253,
final_local_scores=0.48796427, pem=-1.4345005,
similarity=26.452578]
```

Sample: b'539_SOCCER'

```
SOCCER - [1 1 GERMAN 1] FIRST DIVISION SUMMARIES . [2 2 BONN
2] 1996-08-27 Summaries of [3 3 Bundesliga 3] matches played
on Tuesday : [4 4 Borussia Dortmund 4] 3 ( [5 5 Riedle 5] 8th
minute , [6 6 Heinrich 6] 29th , [7 7 Tretsckok 7] 77th ) [8 8
Freiburg 8] 1 ( [9 9 Decheiver 9] 51st penalty ) . Halftime
2-0 . Attendance 48,800 . [10 10 Hamburg 10] 0 [11 11 VfB Stu
ttgart 11] 4 ( [12 12 Balakov 12] 29th , [13 13 Bobic 13]
47th and 60th , [14 14 Hagner 14] 85th ) . 0-1 . 31,139 . [15
15 Werder Bremen 15] 1 ( Schulz 31st ) [16 16 Borussia Moench
engladbach 16] 0 . 1-0 . 24,800 . [17 17 Schalke 17] 1 ( [18 1
8 Thon 18] 2nd ) [19 19 Bochum 19] 1 ( Donkow 86th ) . 1-0 .
33,230 .
1: 11867 Germany
2: 3295 Bonn
3: 686142 Bundesliga
4: 331715 Borussia_Dortmund
5: 2419849 Karl-Heinz_Riedle
6: 5570243 Jörg_Heinrich
7: 9998002 René_Tretsckok
8: 1438835 SC_Freiburg
9: 9383177 Harry_Decheiver
10: 535879 Hamburger_SV
11: 737608 VfB_Stuttgart
12: 1728044 Krasimir_Balakov
13: 2194769 Fredi_Bobic
14: 20883767 Matthias_Hagner
15: 7598082 SV_Werder_Bremen
16: 263466 Borussia_Mönchengladbach
17: 410348 FC_Schalke_04
18: 4909208 Olaf_Thon
19: 513623 VfL_Bochum
1: 11867 Germany
    [pem_pos=1, final_scores=0.7686698,
```

```
final_local_scores=0.8588354, pem=-1.0465823,
similarity=35.410553]
2: 3295 Bonn
    [pem_pos=0, final_scores=0.38592225,
final_local_scores=0.45831525, pem=0.0, similarity=17.766754]
3: 686142 Bundesliga
    [pem_pos=0, final_scores=0.3014539,
final_local_scores=0.23778486, pem=-0.08715299,
similarity=11.659639]
4: 331715 Borussia_Dortmund
    [pem_pos=0, final_scores=0.5428149,
final_local_scores=0.49439853, pem=0.0, similarity=18.843552]
5: 2419849 Karl-Heinz_Riedle
    [pem_pos=0, final_scores=0.64098,
final_local_scores=0.66454715, pem=0.0, similarity=23.921125]
6: 5570243 Jörg_Heinrich
    [pem_pos=1, final_scores=0.7736929,
final_local_scores=0.82118624, pem=-2.0650008,
similarity=39.825344]
7: 9998002 René_Tretschok
    [pem_pos=0, final_scores=0.48049957,
final_local_scores=0.57464665, pem=0.0, similarity=21.238316]
8: 1438835 SC_Freiburg
    [pem_pos=2, final_scores=0.9216805,
final_local_scores=0.932393, pem=-1.7450409,
similarity=41.403984]
9: 9383177 Harry_Decheiver
    [pem_pos=0, final_scores=0.6515075,
final_local_scores=0.7704479, pem=0.0, similarity=27.081415]
10: 535879 Hamburger_SV
    [pem_pos=1, final_scores=0.7639895,
final_local_scores=0.73618567, pem=-2.820485,
similarity=41.397198]
11: 737608 VfB_Stuttgart
    [pem_pos=0, final_scores=0.5516767,
final_local_scores=0.5055498, pem=0.0, similarity=19.176327]
12: 1728044 Krasimir_Balakov
    [pem_pos=0, final_scores=0.49533814,
final_local_scores=0.53983593, pem=0.0, similarity=20.199493]
13: 2194769 Fredi_Bobic
    [pem_pos=0, final_scores=0.5166492,
final_local_scores=0.5653662, pem=0.0, similarity=20.961369]
14: 20883767 Matthias_Hagner
    [pem_pos=5, final_scores=0.4957217,
final_local_scores=0.60562456, pem=-1.7917595,
similarity=31.906626]
15: 7598082 SV_Werder_Bremen
    [pem_pos=0, final_scores=0.60660046,
final_local_scores=0.56664187, pem=0.0, similarity=20.999437]
16: 263466 Borussia_Mönchengladbach
    [pem_pos=0, final_scores=0.49749225,
final_local_scores=0.44711608, pem=0.0, similarity=17.432549]
```

```
17: 410348 FC_Schalke_04
    [pem_pos=0, final_scores=0.50761396,
final_local_scores=0.45893806, pem=0.0, similarity=17.78534]
18: 4909208 Olaf_Thon
    [pem_pos=0, final_scores=0.5619973,
final_local_scores=0.60456055, pem=-0.3051171,
similarity=23.790281]
19: 513623 VfL_Bochum
    [pem_pos=1, final_scores=0.59214,
final_local_scores=0.5620212, pem=-1.509258,
similarity=29.069126]
```

Sample: b'540_SOCCER'

```
SOCCER - [1 1 GERMAN 1] FIRST DIVISION RESULTS . [2 2 BONN 2]
1996-08-27 Results of [3 3 Bundesliga 3] matches played on
Tuesday : [4 4 Borussia Dortmund 4] 3 [5 5 Freiburg 5] 1 [6 6
Hamburg 6] 0 [7 7 VfB Stuttgart 7] 4 [8 8 Werder Bremen 8] 1 [
9 9 Borussia Moenchengladbach 9] 0 [10 10 Schalke 10] 1 [11 11
Bochum 11] 1 [12 12 Bundesliga 12] standings after Tuesday 's
games ( tabulate under played , won , drawn , lost , goals
for , goals against , points ) : [13 13 VfB Stuttgart 13] 3 3
0 0 10 1 9 [14 14 Borussia Dortmund 14] 4 3 0 1 12 6 9 [15 15
Cologne 15] 3 3 0 0 7 1 9 [16 16 Bayern Munich 16] 3 2 1 0 7
2 7 [17 17 Bayer Leverkusen 17] 3 2 0 1 7 4 6 [18 18 VfL Boch
um 18] 4 1 3 0 4 3 6 [19 19 Hamburg 19] 4 2 0 2 7 7 6 [20 20 K
arlsruhe 20] 2 1 1 0 5 3 4 [21 21 St Pauli 21] 3 1 1 1 7 7 4 [
22 22 Werder Bremen 22] 4 1 1 2 5 6 4 [23 23 1860 Munich 23]
3 1 0 2 3 5 3 [24 24 Schalke 24] 4 0 3 1 5 9 3 [25 25 Fortuna
Duesseldorf 25] 3 1 0 2 1 7 3 [26 26 Freiburg 26] 4 1 0 3 6
13 3 [27 27 Hansa Rostock 27] 3 0 2 1 3 4 2 [28 28 Arminia Bi
elefeld 28] 3 0 2 1 2 3 2 [29 29 Borussia Moenchengladbach 29]
4 0 2 2 1 4 2 [30 30 MSV Duisburg 30] 3 0 0 3 1 8 0
```

```
1: 11867 Germany
2: 3295 Bonn
3: 686142 Bundesliga
4: 331715 Borussia_Dortmund
5: 1438835 SC_Freiburg
6: 535879 Hamburger_SV
7: 737608 VfB_Stuttgart
8: 7598082 SV_Werder_Bremen
9: 263466 Borussia_Moenchengladbach
10: 410348 FC_Schalke_04
11: 513623 VfL_Bochum
12: 686142 Bundesliga
13: 737608 VfB_Stuttgart
14: 331715 Borussia_Dortmund
15: 497888 1._FC_Koeln
16: 172326 FC_Bayern_Munich
17: 502005 Bayer_04_Leverkusen
18: 513623 VfL_Bochum
19: 535879 Hamburger_SV
20: 1734592 Karlsruher_SC
```

```
21: 881244 FC_St._Pauli
22: 7598082 SV_Werder_Bremen
23: 1007441 TSV_1860_München
24: 410348 FC_Schalke_04
25: 1949688 Fortuna_Düsseldorf
26: 1438835 SC_Freiburg
27: 599978 F.C._Hansa_Rostock
28: 1145228 Arminia_Bielefeld
29: 263466 Borussia_Mönchengladbach
30: 1586579 MSV_Duisburg
1: 11867 Germany
    [pem_pos=1, final_scores=0.7825478,
final_local_scores=0.87668866, pem=-1.0465823,
similarity=35.94333]
2: 3295 Bonn
    [pem_pos=0, final_scores=0.37772477,
final_local_scores=0.45504123, pem=0.0, similarity=17.66905]
3: 686142 Bundesliga
    [pem_pos=0, final_scores=0.37152892,
final_local_scores=0.30032682, pem=-0.08715299,
similarity=13.526016]
4: 331715 Borussia_Dortmund
    [pem_pos=0, final_scores=0.5367337,
final_local_scores=0.4785546, pem=0.0, similarity=18.370737]
5: 1438835 SC_Freiburg
    [pem_pos=2, final_scores=0.9491722,
final_local_scores=0.95149785, pem=-1.7450409,
similarity=41.97411]
6: 535879 Hamburger_SV
    [pem_pos=1, final_scores=0.916364,
final_local_scores=0.8915649, pem=-2.820485,
similarity=46.034027]
7: 737608 VfB_Stuttgart
    [pem_pos=0, final_scores=0.5056154,
final_local_scores=0.44720328, pem=0.0, similarity=17.43515]
8: 7598082 SV_Werder_Bremen
    [pem_pos=0, final_scores=0.62412614,
final_local_scores=0.5753389, pem=0.0, similarity=21.258974]
9: 263466 Borussia_Mönchengladbach
    [pem_pos=0, final_scores=0.52419335,
final_local_scores=0.46631056, pem=0.0, similarity=18.00535]
10: 410348 FC_Schalke_04
    [pem_pos=0, final_scores=0.5401158,
final_local_scores=0.48103106, pem=0.0, similarity=18.44464]
11: 513623 VfL_Bochum
    [pem_pos=1, final_scores=0.52233243,
final_local_scores=0.47153723, pem=-1.509258,
similarity=26.368904]
12: 686142 Bundesliga
    [pem_pos=0, final_scores=0.37115127,
final_local_scores=0.29991427, pem=-0.08715299,
similarity=13.513705]
```

13: 737608 VfB_Stuttgart
[pem_pos=0, final_scores=0.47888345,
final_local_scores=0.41800272, pem=0.0, similarity=16.563747]
14: 331715 Borussia_Dortmund
[pem_pos=0, final_scores=0.47232264,
final_local_scores=0.40819532, pem=0.0, similarity=16.271076]
15: 497888 1._FC_Köln
[pem_pos=10, final_scores=0.7316286,
final_local_scores=0.6927445, pem=-3.09805,
similarity=41.610268]
16: 172326 FC_Bayern_Munich
[pem_pos=0, final_scores=0.4521563,
final_local_scores=0.3904366, pem=0.0, similarity=15.741119]
17: 502005 Bayer_04_Leverkusen
[pem_pos=0, final_scores=0.53430384,
final_local_scores=0.4773425, pem=0.0, similarity=18.334564]
18: 513623 VfL_Bochum
[pem_pos=0, final_scores=0.5598559,
final_local_scores=0.5125259, pem=0.0, similarity=19.384508]
19: 535879 Hamburger_SV
[pem_pos=1, final_scores=0.8856979,
final_local_scores=0.8580668, pem=-2.820485,
similarity=45.034374]
20: 1734592 Karlsruher_SC
[pem_pos=1, final_scores=0.6775114,
final_local_scores=0.6410455, pem=-2.2263577,
similarity=35.32707]
21: 881244 FC_St._Pauli
[pem_pos=1, final_scores=0.55388564,
final_local_scores=0.518991, pem=-0.11686432,
similarity=20.212965]
22: 7598082 SV_Werder_Bremen
[pem_pos=0, final_scores=0.6172531,
final_local_scores=0.56783116, pem=0.0, similarity=21.034927]
23: 1007441 TSV_1860_München
[pem_pos=0, final_scores=0.46139884,
final_local_scores=0.40843993, pem=0.0, similarity=16.278376]
24: 410348 FC_Schalke_04
[pem_pos=0, final_scores=0.5389797,
final_local_scores=0.4797901, pem=0.0, similarity=18.407606]
25: 1949688 Fortuna_Düsseldorf
[pem_pos=0, final_scores=0.63537234,
final_local_scores=0.5986173, pem=0.0, similarity=21.953648]
26: 1438835 SC_Freiburg
[pem_pos=2, final_scores=0.87438256,
final_local_scores=0.86980146, pem=-1.7450409,
similarity=39.536125]
27: 599978 F.C._Hansa_Rostock
[pem_pos=0, final_scores=0.54117376,
final_local_scores=0.50219166, pem=0.0, similarity=19.076115]
28: 1145228 Arminia_Bielefeld
[pem_pos=0, final_scores=0.5133228,

```
final_local_scores=0.46483648, pem=0.0, similarity=17.96136]
29: 263466 Borussia_Mönchengladbach
    [pem_pos=0, final_scores=0.5332466,
final_local_scores=0.47619992, pem=0.0, similarity=18.300468]
30: 1586579 MSV_Duisburg
    [pem_pos=0, final_scores=0.5255483,
final_local_scores=0.48045933, pem=0.0, similarity=18.427578]
```

Sample: b'541_SOCCER'

```
SOCCER - [1 1 DUTCH 1] FIRST DIVISION SUMMARY . [2 2 AMSTERDAM
AM 2] 1996-08-27 [3 3 Dutch 3] first division summary on
Tuesday : [4 4 Fortuna Sittard 4] 2 ( 1 Jeffrey 7 , Roest
33 ) [5 5 Heerenveen 5] 4 ( [6 6 Korneev 6] 15 , Hansma 24 ,
Wouden 70 , 90 ) . Halftime 2-2 .
1: 21148 Netherlands
2: 844 Amsterdam
3: 21148 Netherlands
4: 1336325 Fortuna_Sittard
5: 745757 SC_Heerenveen
6: 3989180 Igor_Korneev
1: 5109955 Michael_Jeffrey
    [winner: 40087376 Jared_Jeffrey, pem_pos=13,
final_scores=-0.44581595, final_local_scores=-0.42802823,
pem=-4.1037445, similarity=13.633297 |
     RECALL MISS: 5109955 Michael_Jeffrey]
1: 21148 Netherlands
    [pem_pos=0, final_scores=0.6788486,
final_local_scores=0.6758274, pem=-0.7943452,
similarity=28.577522]
2: 844 Amsterdam
    [pem_pos=0, final_scores=0.6123728,
final_local_scores=0.62530875, pem=-0.0642803,
similarity=23.099741]
3: 21148 Netherlands
    [pem_pos=0, final_scores=0.5348327,
final_local_scores=0.5185119, pem=-0.7943452,
similarity=23.882915]
4: 1336325 Fortuna_Sittard
    [pem_pos=0, final_scores=0.69954306,
final_local_scores=0.74356437, pem=0.0, similarity=26.279158]
5: 745757 SC_Heerenveen
    [pem_pos=1, final_scores=0.58026,
final_local_scores=0.6146812, pem=-0.6351416,
similarity=25.887026]
6: 3989180 Igor_Korneev
    [pem_pos=0, final_scores=0.5268956,
final_local_scores=0.642503, pem=0.0, similarity=23.263285]
```

Sample: b'542_SOCCER'

```
SOCCER - [1 1 DUTCH 1] FIRST DIVISION RESULT . [2 2 AMSTERDAM
2] 1996-08-27 Result of a [3 3 Dutch 3] first division soccer
match played on Tuesday : [4 4 Fortuna Sittard 4] 2 [5 5 Heer
```

```
enveen 5] 4
1: 21148 Netherlands
2: 844 Amsterdam
3: 21148 Netherlands
4: 1336325 Fortuna_Sittard
5: 745757 SC_Herenveen
1: 21148 Netherlands
    [pem_pos=0, final_scores=0.6803512,
final_local_scores=0.6651399, pem=-0.7943452,
similarity=28.258587]
2: 844 Amsterdam
    [pem_pos=0, final_scores=0.6298213,
final_local_scores=0.63262385, pem=-0.0642803,
similarity=23.318037]
3: 21148 Netherlands
    [pem_pos=0, final_scores=0.6844498,
final_local_scores=0.669617, pem=-0.7943452,
similarity=28.392193]
4: 1336325 Fortuna_Sittard
    [pem_pos=0, final_scores=0.6487897,
final_local_scores=0.6954457, pem=0.0, similarity=24.8432]
5: 745757 SC_Herenveen
    [pem_pos=1, final_scores=0.5691801,
final_local_scores=0.61578184, pem=-0.6351416,
similarity=25.91987]
```

Sample: b'543_ICE'

ICE HOCKEY - [1 1 FINLAND 1] BEAT [2 2 CZECH REPUBLIC 2] IN WORLD CUP MATCH . [3 3 HELSINKI 3] 1996-08-27 [4 4 Finland 4] beat the [5 5 Czech Republic 5] 7-3 (period scores 4-1 1-1 2-1) in their ice hockey [6 6 World Cup 6] , [7 7 European 7] group match on Tuesday . Scorers : [8 8 Finland 8] - [9 9 Ville Peltonen 9] (10th minute) , [10 10 Juha Ylonen 10] (10th) , [11 11 Teemu Selanne 11] (11th) , [12 12 Jyrki Lumme 12] (13th and 51st) , [13 13 Janne Ojanen 13] (23rd) , [14 14 Christian Ruutu 14] (45th) [15 15 Czech Republic 15] - [16 16 Radek Bonk 16] (7th) , [17 17 Robert Reichel 17] (33rd , penalty) , [18 18 Jiri Dopita 18] (57th) Standings (tabulate under played , won , drawn , lost , goals for , against , points) : [19 19 Sweden 19] 1 1 0 0 6 1 2 [20 20 Finland 20] 1 1 0 0 7 3 2 [21 21 Czech Republic 21] 1 0 0 1 3 7 0 [22 22 Germany 22] 1 0 0 1 1 6 0

1: 950383 Finland_men's_national_ice_hockey_team
2: 949898 Czech_Republic_men's_national_ice_hockey_team
3: 13696 Helsinki
4: 950383 Finland_men's_national_ice_hockey_team
5: 949898 Czech_Republic_men's_national_ice_hockey_team
6: 263861 World_Cup_of_Hockey
7: 9239 Europe
8: 950383 Finland_men's_national_ice_hockey_team
9: 3078614 Ville_Peltonen
10: 6277048 Juha_Ylönen

```
11: 264865 Teemu_Selänne
12: 1182657 Jyrki_Lumme
13: 11299099 Janne_Ojanen
14: 4570035 Christian_Ruuttu
15: 949898 Czech_Republic_men's_national_ice_hockey_team
16: 569462 Radek_Bonk
17: 3651889 Robert_Reichel
18: 4077824 Jiří_Dopita
19: 747522 Sweden_men's_national_ice_hockey_team
20: 950383 Finland_men's_national_ice_hockey_team
21: 949898 Czech_Republic_men's_national_ice_hockey_team
22: 952624 Germany_men's_national_ice_hockey_team
1: 950383 Finland_men's_national_ice_hockey_team
    [pem_pos=2, final_scores=0.36185622,
final_local_scores=0.30050147, pem=-3.8760931,
similarity=34.136074]
2: 949898 Czech_Republic_men's_national_ice_hockey_team
    [pem_pos=2, final_scores=0.45568806,
final_local_scores=0.4131379, pem=-3.5565252,
similarity=35.75951]
3: 13696 Helsinki
    [pem_pos=0, final_scores=0.5255977,
final_local_scores=0.59096265, pem=-0.015504186,
similarity=21.809532]
4: 950383 Finland_men's_national_ice_hockey_team
    [pem_pos=2, final_scores=0.48622078,
final_local_scores=0.4361875, pem=-3.8760931,
similarity=38.185215]
5: 949898 Czech_Republic_men's_national_ice_hockey_team
    [pem_pos=2, final_scores=0.46406192,
final_local_scores=0.4217311, pem=-3.5565252,
similarity=36.01595]
6: 263861 World_Cup_of_Hockey
    [pem_pos=16, final_scores=0.24119602,
final_local_scores=0.23301476, pem=-4.109184,
similarity=33.389717]
7: 9239 Europe
    [pem_pos=0, final_scores=0.33802986,
final_local_scores=0.46148425, pem=-0.25557294,
similarity=19.25117]
8: 950383 Finland_men's_national_ice_hockey_team
    [pem_pos=2, final_scores=0.5815827,
final_local_scores=0.5403559, pem=-3.8760931,
similarity=41.293808]
9: 3078614 Ville_Peltonen
    [pem_pos=0, final_scores=0.68105096,
final_local_scores=0.7202416, pem=0.0, similarity=25.583158]
10: 6277048 Juha_Ylönen
    [pem_pos=0, final_scores=0.55380493,
final_local_scores=0.6195221, pem=0.0, similarity=22.577488]
11: 264865 Teemu_Selänne
    [pem_pos=0, final_scores=0.7638598,
```

```
final_local_scores=0.7906107, pem=0.0, similarity=27.683113]
12: 1182657 Jyrki_Lumme
    [pem_pos=0, final_scores=0.5303031,
final_local_scores=0.5708197, pem=0.0, similarity=21.124111]
13: 11299099 Janne_Ojanen
    [pem_pos=0, final_scores=0.5237001,
final_local_scores=0.56909835, pem=0.0, similarity=21.072742]
14: 4570035 Christian_Ruutu
    [pem_pos=0, final_scores=0.4506986,
final_local_scores=0.5390037, pem=0.0, similarity=20.174658]
15: 949898 Czech_Republic_men's_national_ice_hockey_team
    [pem_pos=2, final_scores=0.47122735,
final_local_scores=0.42955828, pem=-3.5565252,
similarity=36.249527]
16: 569462 Radek_Bonk
    [pem_pos=0, final_scores=0.50140977,
final_local_scores=0.57363814, pem=0.0, similarity=21.20822]
17: 3651889 Robert_Reichel
    [pem_pos=0, final_scores=0.7202284,
final_local_scores=0.75100297, pem=0.0, similarity=26.501139]
18: 4077824 Jiří_Dopita
    [pem_pos=0, final_scores=0.64808315,
final_local_scores=0.7000091, pem=0.0, similarity=24.979382]
19: 747522 Sweden_men's_national_ice_hockey_team
    [pem_pos=2, final_scores=0.35646212,
final_local_scores=0.30889827, pem=-4.1310015,
similarity=35.77288]
20: 950383 Finland_men's_national_ice_hockey_team
    [pem_pos=2, final_scores=0.5323569,
final_local_scores=0.4865842, pem=-3.8760931,
similarity=39.68915]
21: 949898 Czech_Republic_men's_national_ice_hockey_team
    [pem_pos=2, final_scores=0.47152585,
final_local_scores=0.4298843, pem=-3.5565252,
similarity=36.259254]
22: 952624 Germany_men's_national_ice_hockey_team
    [pem_pos=9, final_scores=0.16298507,
final_local_scores=0.12790447, pem=-4.981943,
similarity=34.999218]
```

Sample: b'544_SOCCER'

SOCCER - [1 1 NEUCHATEL 1] TO APPEAL AGAINST [2 2 CYPRIEN 2]
'S NINE-MONTH BAN . [3 3 GENEVA 3] 1996-08-27 [4 4 Swiss 4]
league leaders [5 5 Neuchatel Xamax 5] said on Tuesday they
would appeal against a nine-month ban imposed on [6 6 French 6]
] international defender [7 7 Jean-Pierre Cyprien 7] for his
part in a post-match brawl . [8 8 Cyprien 8] , also fined
10,000 [9 9 Swiss 9] francs (\$ 8,400) , traded punches with
[10 10 St Gallen 10] 's [11 11 Brazilian 11] player Claudio
Moura after a match on Saturday . When officials and coaching
staff tried to intervene , [12 12 Cyprien 12] launched a
flying kick at Moura , but only succeeded in kneeing [13 13 St

Gallen 13] coach Roger Hegi in the stomach . Moura , who appeared to have elbowed [14 14 Cyprien 14] in the final minutes of the 3-0 win by [15 15 Neuchatel 15] , was suspended for seven matches and fined 1,000 francs (\$ 840) by the [16 16 Swiss 16] league disciplinary committee . Club president Gilbert Facchinetti said he was astonished the committee had arrived at its decision so quickly and vowed the club would appeal . [17 17 Neuchatel 17] coach [18 18 Gilbert Gress 18] described the incident as " shocking " , but said Moura was also to blame . " Moura physically and verbally provoked [19 19 Cyprien 19] during the match . The referee could not have seen it or he would have punished him , " [20 20 Gress 20] said . " During the scuffle , Moura threw the first punch . Tomorrow , if someone punches me , I would not know how to react . " [21 21 Cyprien 21] , who won his one [22 22 French 22] cap against [23 23 Italy 23] in February 1994 , cannot play in [24 24 Switzerland 24] or elsewhere until May next year .

1: 494581 Neuchâtel
2: 11942980 Jean-Pierre_Cyprien
3: 12521 Geneva
4: 26748 Switzerland
5: 2228737 Neuchâtel_Xamax
6: 5843419 France
7: 11942980 Jean-Pierre_Cyprien
8: 11942980 Jean-Pierre_Cyprien
9: 26748 Switzerland
10: 403902 St. Gallen
11: 3383 Brazil
12: 11942980 Jean-Pierre_Cyprien
13: 1607312 FC_St._Gallen
14: 11942980 Jean-Pierre_Cyprien
15: 494581 Neuchâtel
16: 26748 Switzerland
17: 494581 Neuchâtel
18: 10533315 Gilbert_Gress
19: 11942980 Jean-Pierre_Cyprien
20: 10533315 Gilbert_Gress
21: 11942980 Jean-Pierre_Cyprien
22: 5843419 France
23: 14532 Italy
24: 26748 Switzerland
1: 494581 Neuchâtel
[pem_pos=0, final_scores=0.17647378, final_local_scores=0.16779685, pem=0.0, similarity=9.097106]
2: 11942980 Jean-Pierre_Cyprien
[pem_pos=0, final_scores=0.67853695, final_local_scores=0.687828, pem=0.0, similarity=24.615873]
3: 12521 Geneva
[pem_pos=0, final_scores=0.46068215, final_local_scores=0.4892671, pem=-0.24772356, similarity=20.037579]

4: 26748 Switzerland
[pem_pos=0, final_scores=0.5104614,
final_local_scores=0.51691264, pem=-0.1392413,
similarity=20.272633]
5: 2228737 Neuchâtel_Xamax
[pem_pos=0, final_scores=0.44407177,
final_local_scores=0.5022934, pem=0.0, similarity=19.07915]
6: 5843419 France
[pem_pos=0, final_scores=0.68630004,
final_local_scores=0.7070281, pem=-0.80223274,
similarity=29.551508]
7: 11942980 Jean-Pierre_Cyprien
[pem_pos=0, final_scores=0.788142,
final_local_scores=0.8017509, pem=0.0, similarity=28.015558]
8: 11942980 Jean-Pierre_Cyprien
[pem_pos=0, final_scores=1.0158511,
final_local_scores=1.0562925, pem=0.0, similarity=35.61159]
9: 26748 Switzerland
[pem_pos=0, final_scores=0.50164455,
final_local_scores=0.5072816, pem=-0.1392413,
similarity=19.985224]
10: 403902 St._Gallen
[pem_pos=0, final_scores=0.3919819,
final_local_scores=0.43669474, pem=0.0, similarity=17.121555]
11: 3383 Brazil
[pem_pos=0, final_scores=0.7609176,
final_local_scores=0.868102, pem=-0.2733083,
similarity=31.481901]
12: 11942980 Jean-Pierre_Cyprien
[pem_pos=0, final_scores=1.048371,
final_local_scores=1.0918155, pem=0.0, similarity=36.67166]
13: 1607312 FC_St._Gallen
[pem_pos=1, final_scores=0.36456925,
final_local_scores=0.42885494, pem=-0.54538625,
similarity=19.853495]
14: 11942980 Jean-Pierre_Cyprien
[pem_pos=0, final_scores=0.86497825,
final_local_scores=0.8914868, pem=0.0, similarity=30.693459]
15: 494581 Neuchâtel
[pem_pos=0, final_scores=0.35870004,
final_local_scores=0.36685133, pem=0.0, similarity=15.037289]
16: 26748 Switzerland
[pem_pos=0, final_scores=0.6736786,
final_local_scores=0.6952026, pem=-0.1392413,
similarity=25.59316]
17: 494581 Neuchâtel
[pem_pos=0, final_scores=0.40681547,
final_local_scores=0.41941017, pem=0.0, similarity=16.60575]
18: 10533315 Gilbert_Gress
[pem_pos=0, final_scores=0.74257046,
final_local_scores=0.86648756, pem=0.0, similarity=29.947432]
19: 11942980 Jean-Pierre_Cyprien

```
[pem_pos=0, final_scores=0.9748228,
final_local_scores=1.0114753, pem=0.0, similarity=34.274155]
20: 10533315 Gilbert_Gress
    [pem_pos=0, final_scores=0.630594,
final_local_scores=0.75679016, pem=0.0, similarity=26.673841]
21: 11942980 Jean-Pierre_Cyprien
    [pem_pos=0, final_scores=0.94000703,
final_local_scores=0.9734443, pem=0.0, similarity=33.139233]
22: 5843419 France
    [pem_pos=0, final_scores=0.69204587,
final_local_scores=0.7133045, pem=-0.80223274,
similarity=29.73881]
23: 14532 Italy
    [pem_pos=0, final_scores=0.6094243,
final_local_scores=0.6607269, pem=-0.20612453,
similarity=24.928059]
24: 26748 Switzerland
    [pem_pos=0, final_scores=0.54371744,
final_local_scores=0.5532399, pem=-0.18598008,
similarity=21.610882]
```

Sample: b'545_CYCLING'

CYCLING - [1 1 BUGNO 1] CLEARED OF DOPING . [2 2 MILAN 2]
1996-08-27 Veteran [3 3 Italian 3] [4 4 Gianni Bugno 4] has
been cleared of doping after testing positive for high levels
of testosterone during the [5 5 Tour of Switzerland 5] in
June , the [6 6 Italian 6] cycling federation said on
Tuesday . " He has been cleared . The case is closed , " a
spokesman said . [7 7 Bugno 7] tested positive for the banned
hormone after the fifth stage of the Tour , in which he
finished third overall . But the spokesman said subsequent
tests in [8 8 Cologne 8] proved his body produced higher-than-
average testosterone levels naturally . [9 9 Bugno 9] , who
won the Giro d'Italia in 1990 and two successive world
titles , was banned for three months in 1994 after testing
positive for the stimulant caffeine .

```
1: 30862327 Gianni_Bugno
2: 36511 Milan
3: 14532 Italy
4: 30862327 Gianni_Bugno
5: 1988710 Tour_de_Suisse
6: 14532 Italy
7: 30862327 Gianni_Bugno
8: 6187 Cologne
9: 30862327 Gianni_Bugno
1: 30862327 Gianni_Bugno
    [pem_pos=1, final_scores=0.63323444,
final_local_scores=0.60181254, pem=0.0, similarity=22.049]
2: 36511 Milan
    [pem_pos=0, final_scores=0.47018707,
final_local_scores=0.5162186, pem=-0.3015942,
similarity=21.134823]
```

```

3: 14532 Italy
    [pem_pos=0, final_scores=0.7136498,
final_local_scores=0.77998596, pem=-0.75948656,
similarity=31.496254]
4: 30862327 Gianni_Bugno
    [pem_pos=0, final_scores=0.8874002,
final_local_scores=0.8718657, pem=0.0, similarity=30.107924]
5: 1988710 Tour_de_Suisse
    [pem_pos=0, final_scores=0.49138707,
final_local_scores=0.56102675, pem=0.0, similarity=20.831871]
6: 14532 Italy
    [pem_pos=0, final_scores=0.58067507,
final_local_scores=0.63473123, pem=-0.75948656,
similarity=27.161564]
7: 30862327 Gianni_Bugno
    [pem_pos=1, final_scores=0.6766258,
final_local_scores=0.64921105, pem=0.0, similarity=23.463465]
8: 6187 Cologne
    [pem_pos=0, final_scores=0.43993014,
final_local_scores=0.533799, pem=-0.08239529,
similarity=20.467417]
9: 30862327 Gianni_Bugno
    [pem_pos=1, final_scores=0.73536474,
final_local_scores=0.7133745, pem=0.0, similarity=25.378231]

```

Sample: b'546_CYCLING'

CYCLING - COLONNA WINS FIRST STAGE OF TOUR OF NETHERLANDS . [1 1 HAARLEM 1] , [2 2 Netherlands 2] 1996-08-27 Leading results and overall standings after the 161 kilometre first stage of the [3 3 Tour of the Netherlands 3] between [4 4 Gouda 4] and [5 5 Haarlem 5] on Tuesday . 1. Federico Colonna ([6 6 Italy 6]) Mapei three hours 43 mins five secs 2. [7 7 Robbie McEwen 7] ([8 8 Australia 8]) [9 9 Rabobank 9] 3. [10 10 Jans Koerts 10] ([11 11 Netherlands 11]) Palmans 4. Sven Teutenberg ([12 12 Germany 12]) US Postal 5. [13 13 Tom Steels 13] ([14 14 Belgium 14]) Mapei 6. [15 15 Endrio Leoni 15] ([16 16 Italy 16]) Aki 7. [17 17 Johan Capiot 17] ([18 18 Belgium 18]) [19 19 Collstrop 19] 8. John den Braber (Neths) [20 20 Collstrop 20] 9. [21 21 Jeroen Blijlevens 21] (Neths) [22 22 TVM 22] 10. Michael van der Wolf (Neths) Foreldorado all same time . Leading overall standings after first stage . 1. Colonna three hours 42 mins 55 seconds 2. [23 23 McEwen 23] 0:04 seconds behind 3. [24 24 Koerts 24] 0:06 4. Gianluca Corini ([25 25 Italy 25]) Aki 0:07 5. Wim Omloop ([26 26 Belgium 26]) [27 27 Collstrop 27] same time 6. [28 28 Lance Armstrong 28] ([29 29 USA 29]) [30 30 Motorola 30] 0:08 7. [31 31 Tristan Hoffman 31] (Neths) [32 32 TVM 32] same time 8. [33 33 George Hincapie 33] ([34 34 USA 34]) [35 35 Motorola 35] 0:09 9. [36 36 John Talen 36] (Neths) Foreldorado same time 10. [37 37 Teutenberg 37] 0:10

1: 63278 Haarlem
2: 21148 Netherlands

```
3: 11065220 Ronde_van_Nederland
4: 102081 Gouda
5: 63278 Haarlem
6: 14532 Italy
7: 1484348 Robbie_McEwen
8: 4689264 Australia
9: 2354465 Belkin_Pro_Cycling_Team
10: 17022543 Jans_Koerts
11: 21148 Netherlands
12: 11867 Germany
13: 10612540 Tom_Steels
14: 3343 Belgium
15: 11099260 Endrio_Leoni
16: 14532 Italy
17: 13573637 Johan_Capiot
18: 3343 Belgium
19: 6932095 Cycle_Collstrop
20: 6932095 Cycle_Collstrop
21: 11066258 Jeroen_Blijlevens
22: 2642474 TVM_(cycling_team)
23: 1484348 Robbie_McEwen
24: 17022543 Jans_Koerts
25: 14532 Italy
26: 3343 Belgium
27: 6932095 Cycle_Collstrop
28: 23243880 Lance_Armstrong
29: 3434750 United_States
30: 2290911 7-Eleven_(cycling_team)
31: 13153870 Tristan_Hoffman
32: 2642474 TVM_(cycling_team)
33: 1257631 George_Hincapie
34: 3434750 United_States
35: 2290911 7-Eleven_(cycling_team)
36: 11016419 John_Talen
37: 10922103 Ina-Yoko_Teutenberg
1: 63278 Haarlem
    [pem_pos=0, final_scores=0.5125707,
final_local_scores=0.5595565, pem=0.0, similarity=20.787994]
2: 21148 Netherlands
    [pem_pos=2, final_scores=0.5302717,
final_local_scores=0.53094727, pem=-0.24940795,
similarity=21.290558]
3: 11065220 Ronde_van_Nederland
    [pem_pos=0, final_scores=0.62023175,
final_local_scores=0.70123893, pem=0.0, similarity=25.01608]
4: 102081 Gouda
    [pem_pos=0, final_scores=0.4321733,
final_local_scores=0.49061418, pem=-0.32439408,
similarity=20.494724]
5: 63278 Haarlem
    [pem_pos=0, final_scores=0.52256113,
final_local_scores=0.57046956, pem=0.0, similarity=21.113663]
```

```
6: 14532 Italy
    [pem_pos=0, final_scores=0.7441847,
final_local_scores=0.8028279, pem=-0.20612453,
similarity=29.168636]
7: 1484348 Robbie_McEwen
    [pem_pos=0, final_scores=0.72238666,
final_local_scores=0.78017366, pem=0.0, similarity=27.37165]
8: 4689264 Australia
    [pem_pos=0, final_scores=0.67383283,
final_local_scores=0.770783, pem=-0.17168477,
similarity=28.025064]
9: 2354465 Belkin_Pro_Cycling_Team
    [pem_pos=0, final_scores=0.65076405,
final_local_scores=0.6963405, pem=0.0, similarity=24.869902]
10: 17022543 Jans_Koerts
    [pem_pos=0, final_scores=0.6486077,
final_local_scores=0.77020365, pem=0.0, similarity=27.074125]
11: 21148 Netherlands
    [pem_pos=2, final_scores=0.6358214,
final_local_scores=0.6462443, pem=-0.24940795,
similarity=24.73125]
12: 11867 Germany
    [pem_pos=0, final_scores=0.8036888,
final_local_scores=0.85717314, pem=-0.15529114,
similarity=30.513966]
13: 10612540 Tom_Steels
    [pem_pos=0, final_scores=0.63101536,
final_local_scores=0.7194125, pem=0.0, similarity=25.558416]
14: 3343 Belgium
    [pem_pos=0, final_scores=0.71361834,
final_local_scores=0.72511595, pem=-0.24180195,
similarity=27.043577]
15: 11099260 Endrio_Leoni
    [pem_pos=0, final_scores=0.7463811,
final_local_scores=0.861847, pem=0.0, similarity=29.808947]
16: 14532 Italy
    [pem_pos=0, final_scores=0.74477714,
final_local_scores=0.80347514, pem=-0.20612453,
similarity=29.187952]
17: 13573637 Johan_Capiot
    [pem_pos=0, final_scores=0.63345194,
final_local_scores=0.76715755, pem=0.0, similarity=26.983225]
18: 3343 Belgium
    [pem_pos=0, final_scores=0.7301199,
final_local_scores=0.7431415, pem=-0.24180195,
similarity=27.581493]
19: 6932095 Cycle_Collstrop
    [pem_pos=0, final_scores=0.5718779,
final_local_scores=0.6524399, pem=0.0, similarity=23.55982]
20: 6932095 Cycle_Collstrop
    [pem_pos=0, final_scores=0.6655669,
final_local_scores=0.75478095, pem=0.0, similarity=26.613882]
```

```
21: 11066258 Jeroen_Blijlevens
    [pem_pos=0, final_scores=0.90608627,
final_local_scores=1.0099921, pem=0.0, similarity=34.229893]
22: 2642474 TVM_(cycling_team)
    [pem_pos=1, final_scores=0.7299539,
final_local_scores=0.8538557, pem=-0.84397006,
similarity=34.16011]
23: 1484348 Robbie_McEwen
    [pem_pos=0, final_scores=1.2445945,
final_local_scores=1.359143, pem=0.0, similarity=44.649254]
24: 17022543 Jans_Koerts
    [pem_pos=0, final_scores=0.5964071,
final_local_scores=0.71318245, pem=0.0, similarity=25.3725]
25: 14532 Italy
    [pem_pos=0, final_scores=0.709997,
final_local_scores=0.765483, pem=-0.20612453,
similarity=28.054192]
26: 3343 Belgium
    [pem_pos=0, final_scores=0.6800758,
final_local_scores=0.68847585, pem=-0.24180195,
similarity=25.950163]
27: 6932095 Cycle_Collstrop
    [pem_pos=0, final_scores=0.5471371,
final_local_scores=0.6254143, pem=0.0, similarity=22.753323]
28: 23243880 Lance_Armstrong
    [pem_pos=0, final_scores=0.59163266,
final_local_scores=0.6339404, pem=0.0, similarity=23.00776]
29: 3434750 United_States
    [pem_pos=0, final_scores=0.5112062,
final_local_scores=0.5705752, pem=-0.13038418,
similarity=21.825865]
30: 2290911 7-Eleven_(cycling_team)
    [pem_pos=1, final_scores=0.72406775,
final_local_scores=0.8106048, pem=-1.7647687,
similarity=37.876865]
31: 13153870 Tristan_Hoffman
    [pem_pos=0, final_scores=0.64423627,
final_local_scores=0.7896012, pem=0.0, similarity=27.652988]
32: 2642474 TVM_(cycling_team)
    [pem_pos=1, final_scores=0.7581282,
final_local_scores=0.88463193, pem=-0.84397006,
similarity=35.078537]
33: 1257631 George_Hincapie
    [pem_pos=0, final_scores=0.83528453,
final_local_scores=0.89819676, pem=0.0, similarity=30.893696]
34: 3434750 United_States
    [pem_pos=0, final_scores=0.52393717,
final_local_scores=0.5844818, pem=-0.13038418,
similarity=22.240866]
35: 2290911 7-Eleven_(cycling_team)
    [pem_pos=1, final_scores=0.71258336,
final_local_scores=0.7980598, pem=-1.7647687,
```

```
similarity=37.502495]
36: 11016419 John_Talen
    [pem_pos=0, final_scores=0.7279259,
final_local_scores=0.8631937, pem=0.0, similarity=29.849134]
37: 10922103 Ina-Yoko_Teutenberg
    [pem_pos=0, final_scores=0.31099677,
final_local_scores=0.40191185, pem=0.0, similarity=16.083565]
```

Sample: b'547_COFINEC'

COFINEC SLIPS ON [1 1 BUDAPEST 1] BOURSE BUT FUTURE STRONG . Emese Bartha [2 2 BUDAPEST 2] 1996-08-27 Expectations that Cofinec S.A. , the [3 3 Hungarian 3] bourse 's first foreign listing , will report a disappointing first half have depressed the stock below its issue price , but analysts expect a rebound in the long term . " The first half of the year is unlikely to be as strong as expected so the company will probably be unable to reach its annual plan in 1996 , " said Gabor Sitanyi , a London-based analyst for ING Barings . The French-registered packaging materials company , which floated its shares in [4 4 Hungary 4] in July , for most of the past two weeks hovered below the 6,425 forints / Global Depository Receipts price of its initial offering , which was oversubscribed . The company , which asked for a two-week delay from the usual August 15 deadline for reporting first-half results , closed on Tuesday at 5,800 forints , down 300 . " Cofinec 's first-half figures will be ... between one-third of two-fifths of its annual plan , " said Tamas Erdei , a [5 5 Budapest-based 5] analyst for ABN-AMRO Hoare Govett . Analysts blame , at least partly , [6 6 Hungary 6] 's macroeconomic environment for the weaker figures for Cofinec which , operating in [7 7 Hungary 7] , [8 8 Poland 8] and the [9 9 Czech Republic 9] , now generates about 55 to 60 percent of its annual sales from [10 10 Hungary 10] . [11 11 Hungary 1 1] 's Gross Domestic Product fell one percentage point in the first quarter while real wages plunged 7.2 percentage points in the first half of 1996 . Both will have their impact on Cofinec 's figures , the analysts said . Despite the current difficulties , however , analysts were convinced that Cofinec 's outlook was strong . " The eastern [12 12 European 12] market offers good chances , " said Erdei . " Just like many other companies on the bourse , Cofinec has big growth opportunities . " " At the same time , it 's an advantage for Cofinec that it has a foreign management which perhaps understands the market better , " Erdei added . " Cofinec is a very good story in the long-term as the per capita packaging consumption is still so low in east [13 13 Europe 13] that a very strong increase can be expected (long-term) , " Sitanyi said , saying that several recent moves by Cofinec boosted its position . Among them , he noted that Cofinec had acquired the outstanding stake in its [14 14 Czech 14] folding company Krpaco a.s. , increasing its ownership to 100 percent , so in the second half the whole of Krpaco 's

figures will be consolidated . The company also repaid some \$ 21 million of debt , well above the originally planned \$ 8 million to \$ 9 million . In addition , its [15 15 Polish 15] operation began with some six weeks of delay due to cold winter weather and the test run was also longer than planned . -- [16 16 Budapest 16] newsroom (36 1) 266 2410

1: 36787 Budapest

2: 36787 Budapest

3: 13275 Hungary

4: 13275 Hungary

5: 36787 Budapest

6: 13275 Hungary

7: 13275 Hungary

8: 22936 Poland

9: 5321 Czech_Republic

10: 13275 Hungary

11: 13275 Hungary

12: 9239 Europe

13: 9239 Europe

14: 5321 Czech_Republic

15: 22936 Poland

16: 36787 Budapest

1: 36787 Budapest

[pem_pos=0, final_scores=0.52981883,
final_local_scores=0.48164165, pem=-0.030229695,
similarity=18.627253]

2: 36787 Budapest

[pem_pos=0, final_scores=0.5887961,
final_local_scores=0.5460654, pem=-0.030229695,
similarity=20.549786]

3: 13275 Hungary

[pem_pos=1, final_scores=0.723691,
final_local_scores=0.67339176, pem=-1.0319586,
similarity=29.797018]

4: 13275 Hungary

[pem_pos=0, final_scores=0.64312196,
final_local_scores=0.5853823, pem=-0.32577527,
similarity=23.330305]

5: 36787 Budapest

[pem_pos=0, final_scores=0.46124625,
final_local_scores=0.40673655, pem=0.0, similarity=16.227543]

6: 13275 Hungary

[pem_pos=0, final_scores=0.73498386,
final_local_scores=0.6857275, pem=-0.32577527,
similarity=26.324806]

7: 13275 Hungary

[pem_pos=0, final_scores=0.61534506,
final_local_scores=0.5550403, pem=-0.32577527,
similarity=22.424839]

8: 22936 Poland

[pem_pos=0, final_scores=0.6349434,
final_local_scores=0.5920915, pem=-0.15333635,

```
similarity=22.592772]
9: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.4328363,
final_local_scores=0.37504178, pem=-0.27794555,
similarity=16.793219]
10: 13275 Hungary
    [pem_pos=0, final_scores=0.7153443,
final_local_scores=0.6642743, pem=-0.32577527,
similarity=25.684599]
11: 13275 Hungary
    [pem_pos=0, final_scores=0.64456314,
final_local_scores=0.58695656, pem=-0.32577527,
similarity=23.377285]
12: 9239 Europe
    [pem_pos=0, final_scores=0.55099356,
final_local_scores=0.56428045, pem=-0.25557294,
similarity=22.318811]
13: 9239 Europe
    [pem_pos=0, final_scores=0.5722015,
final_local_scores=0.5874469, pem=-0.086338006,
similarity=22.08982]
14: 5321 Czech_Republic
    [pem_pos=1, final_scores=0.62217134,
final_local_scores=0.58186156, pem=-1.208016,
similarity=28.023]
15: 22936 Poland
    [pem_pos=0, final_scores=0.604225,
final_local_scores=0.55853635, pem=-0.81518245,
similarity=25.190641]
16: 36787 Budapest
    [pem_pos=0, final_scores=0.62626225,
final_local_scores=0.5869914, pem=-0.030229695,
similarity=21.771101]
```

Sample: b'548_DIRECT'

DIRECT EQUITY TRADES ON THE [1 1 CZECH 1] PSE - AUG 27 . [2 2
PRAGUE 2] 1996-08-27 The following is a list of direct equity
trades made on the [3 3 Prague Stock Exchange 3] : ISSUE
Min . Price Max . Price Volume Turnover (CZK) (CZK)
(shares) (CZK 000 's) AGROTONZ TLUMACOV 336.47 336.47
59440 19999.777 AVIA 290.00 290.00 700 203.000 BARUM HOLDING
171.00 171.00 14432 2467.872 CESKA SPORITELNA 335.00 375.00
533153 198354.941 CKD PRAHA HOLDING 369.66 384.00 5565
2065.260 EMKAM 25.00 25.00 34684 867.100 KABLO KLADNO 960.00
960.00 2230 2140.800 [4 4 KOMERCNI BANKA 4] 2320.00 2370.00
7000 16408.700 LECIVA PRAHA 2470.00 2470.00 1360 3359.200
METROSTAV 3024.95 3024.95 3000 9074.850 MORAV.CHEMIC. ZAV .
637.50 637.50 1626 1036.575 [5 5 OKD 5] 111.50 112.56 95975
10752.092 PF IKS KB PLUS 156.00 156.00 6000 936.000 RIF
900.00 900.00 5500 4950.000 SELIKO 4000.00 20000.00 3565
32607.500 SOKOLOVSKA UHELNA 785.00 785.00 6000 4710.000 SPIF
CESKY 339.00 340.00 7546 2562.094 SPT TELECOM 3355.00 3404.71

10700 36337.137 SKODA PLZEN 1045.56 1060.00 10772 11361.330
TABAK 6700.00 6700.00 1000 6700.000 TRINECKE ZELEZARNY 210.00
210.00 3000 630.000 VODNI STAVBY PRAHA 1915.00 1915.00 2000
3830.000 -- Prague Newsroom , 42-2-2423-0003

1: 5321 Czech_Republic
2: 23844 Prague
3: 2332279 Prague_Stock_Exchange
4: 9815359 Komerční_banká
5: 15878109 OKD
1: 5321 Czech_Republic
[pem_pos=1, final_scores=0.4753527,
final_local_scores=0.4899916, pem=-1.208016,
similarity=25.281418]
2: 23844 Prague
[pem_pos=0, final_scores=0.44757628,
final_local_scores=0.4703576, pem=-0.05392049,
similarity=18.41935]
3: 2332279 Prague_Stock_Exchange
[pem_pos=0, final_scores=0.36808294,
final_local_scores=0.46279782, pem=0.0, similarity=17.900522]
4: 9815359 Komerční_banká
[pem_pos=0, final_scores=0.40548283,
final_local_scores=0.4923131, pem=0.0, similarity=18.781317]
5: 15878109 OKD
[pem_pos=0, final_scores=0.46849793,
final_local_scores=0.5653035, pem=0.0, similarity=20.959497]

Sample: b'549_AFTER'

AFTER THE BELL - After hours slows in light volume . [1 1 NEW YORK 1] 1996-08-27 Traders said on Tuesday after-hours activity was light . Both WorldCom Inc and MFS Communications Co Inc were trading but they moved in line with their close . WorldCom , which said it will buy [2 2 MFS 2] , shed 1-3/4 to close at 21 while [3 3 MFS 3] lost 3-8/16 to close at 41-5/16 . The [4 4 New York Stock Exchange 4] said its session one volume was 5,700 shares compared to 53,400 shares Monday . Session two volume was 4,153,800 shares compared to no volume Monday . The [5 5 American Stock Exchange 5] said there was no after-hours activity .

1: 8210131 New_York
2: 10611734 Metropolitan_Fiber_Systems
3: 10611734 Metropolitan_Fiber_Systems
4: 21560 New_York_Stock_Exchange
5: 1488 American_Stock_Exchange
1: 8210131 New_York
[pem_pos=0, final_scores=0.09412579,
final_local_scores=0.16609988, pem=-0.2889045,
similarity=10.617572]
2: 10611734 Metropolitan_Fiber_Systems
[pem_pos=2, final_scores=0.5959974,
final_local_scores=0.668937, pem=-1.2862492,
similarity=31.04695]

```
3: 10611734 Metropolitan_Fiber_Systems
    [pem_pos=2, final_scores=0.6544141,
final_local_scores=0.73274845, pem=-1.2862492,
similarity=32.95121]
4: 21560 New_York_Stock_Exchange
    [pem_pos=0, final_scores=0.45362097,
final_local_scores=0.45044887, pem=0.0, similarity=17.532005]
5: 1488 American_Stock_Exchange
    [pem_pos=0, final_scores=0.43035406,
final_local_scores=0.45617098, pem=0.0, similarity=17.702765]
```

Sample: b'550_CBOE'

[1 1 CBOE 1] in routine review of [2 2 MFS 2] options . [3 3 CHICAGO 3] 1996-08-27 The [4 4 Chicago Board Options Exchange 4] ([5 5 CBOE 5]) said on Tuesday it was doing a routine investigation into trading in options on MFS Communications Co Inc shares . On Monday , the company said it had agreed to be acquired by WorldCom Inc in a deal valued at \$ 14 billion . [6 6 MFS 6] shares surged on the news while WorldCom fell on fears of dilution . [7 7 The New York Times 7] said on Tuesday some of the options trading in [8 8 MFS 8] last Friday may suggest insider trading . [9 9 MFS 9] options also trade on the [10 10 American Stock Exchange 10] and the [11 11 Pacific Stock Exchange 11] . A spokesman for the [12 12 American Stock Exchange 12] would neither confirm or deny whether the exchange was looking into trading . " If there is unusual activity , certainly we look at it , but that 's not to say we 're doing anything official , " he said . [13 13 Pacific Stock Exchange 13] officials were not available . One trader said trading in [14 14 MFS 14] options had increased steadily from about mid-August , and doubted whether any of last Friday 's activity was insider trading . - Derivatives desk , 312 408-8750 / E-mail : derivatives@reuters.com

```
1: 1351813 Chicago_Board_Options_Exchange
2: 10611734 Metropolitan_Fiber_Systems
3: 6886 Chicago
4: 1351813 Chicago_Board_Options_Exchange
5: 1351813 Chicago_Board_Options_Exchange
6: 10611734 Metropolitan_Fiber_Systems
7: 30680 The_New_York_Times
8: 10611734 Metropolitan_Fiber_Systems
9: 10611734 Metropolitan_Fiber_Systems
10: 1488 American_Stock_Exchange
11: 230088 Pacific_Exchange
12: 1488 American_Stock_Exchange
13: 230088 Pacific_Exchange
14: 10611734 Metropolitan_Fiber_Systems
1: 1351813 Chicago_Board_Options_Exchange
    [pem_pos=0, final_scores=0.59536034,
final_local_scores=0.6230851, pem=0.0, similarity=22.683815]
2: 10611734 Metropolitan_Fiber_Systems
    [pem_pos=2, final_scores=0.6981447,
```

```
final_local_scores=0.74667865, pem=-1.2862492,
similarity=33.366917]
3: 6886 Chicago
    [pem_pos=0, final_scores=0.54583347,
final_local_scores=0.6147732, pem=-0.20658566,
similarity=23.559216]
4: 1351813 Chicago_Board_Options_Exchange
    [pem_pos=0, final_scores=0.58569276,
final_local_scores=0.6125247, pem=0.0, similarity=22.368671]
5: 1351813 Chicago_Board_Options_Exchange
    [pem_pos=0, final_scores=0.42632037,
final_local_scores=0.4384346, pem=0.0, similarity=17.173475]
6: 10611734 Metropolitan_Fiber_Systems
    [pem_pos=2, final_scores=0.6385906,
final_local_scores=0.6816248, pem=-1.2862492,
similarity=31.42558]
7: 30680 The_New_York_Times
    [pem_pos=0, final_scores=0.38876784,
final_local_scores=0.5154137, pem=0.0, similarity=19.470686]
8: 10611734 Metropolitan_Fiber_Systems
    [pem_pos=2, final_scores=0.69778705,
final_local_scores=0.746288, pem=-1.2862492,
similarity=33.35526]
9: 10611734 Metropolitan_Fiber_Systems
    [pem_pos=2, final_scores=0.72833157,
final_local_scores=0.77965325, pem=-1.2862492,
similarity=34.350945]
10: 1488 American_Stock_Exchange
    [pem_pos=0, final_scores=0.564485,
final_local_scores=0.5746311, pem=0.0, similarity=21.23785]
11: 230088 Pacific_Exchange
    [pem_pos=0, final_scores=0.4718855,
final_local_scores=0.50771683, pem=0.0, similarity=19.240995]
12: 1488 American_Stock_Exchange
    [pem_pos=0, final_scores=0.4651975,
final_local_scores=0.46617454, pem=0.0, similarity=18.001291]
13: 230088 Pacific_Exchange
    [pem_pos=0, final_scores=0.50455415,
final_local_scores=0.5434023, pem=0.0, similarity=20.305922]
14: 10611734 Metropolitan_Fiber_Systems
    [pem_pos=2, final_scores=0.62676793,
final_local_scores=0.6687104, pem=-1.2862492,
similarity=31.040186]
```

Sample: b'551_Faulding'

Faulding target of patent lawsuit . [1 1 ELIZABETH 1] , [2 2 N.J. 2] 1996-08-27 Faulding Inc said on Tuesday Purdue Frederick Co filed a patent infringement lawsuit against Faulding and its Purepac Pharamceutical unit . The suit was filed because of Purepac 's manufacture of Kadian , a sustained release morphine product , Faulding said . Faulding said the claims in the lawsuit are without merit and will not

impact upon the launch of Kadian in the [3 3 United States 3] . Kadian was approved for sale in the [4 4 United States 4] last month , Faulding said . Zeneca Group Plc , which will market Kadian , was named in the lawsuit with F.H. Faulding & Co , the majority shareholder of Faulding Inc , the company said .

```
1: 125778 Elizabeth,_New_Jersey
2: 21648 New_Jersey
3: 3434750 United_States
4: 3434750 United_States
1: 125778 Elizabeth,_New_Jersey
    [pem_pos=0, final_scores=0.3952672,
final_local_scores=0.3889609, pem=-1.7086731,
similarity=24.989113]
2: 21648 New_Jersey
    [pem_pos=0, final_scores=0.6774381,
final_local_scores=0.6557502, pem=0.0, similarity=23.658607]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.5455453,
final_local_scores=0.5065123, pem=-0.073457584,
similarity=19.604525]
4: 3434750 United_States
    [pem_pos=0, final_scores=0.53556305,
final_local_scores=0.49560827, pem=-0.073457584,
similarity=19.279127]
```

Sample: b'552_McGrath'

[1 1 McGrath 1] left out of [2 2 Ireland 2] [3 3 World Cup 3] squad . [4 4 DUBLIN 4] 1996-08-27 [5 5 Ireland 5] 's most experienced player , defender [6 6 Paul McGrath 6] , was left out of the national squad for the first time in 11 years on Tuesday when new manager [7 7 Mick McCarthy 7] named his side to face [8 8 Liechtenstein 8] in a World Cup qualifier . The 36-year-old [9 9 Aston Villa 9] player won the last of his [10 10 Irish 10] record of 82 international caps against the [11 1 1 Czech Republic 11] in [12 12 Prague 12] in April . " Paul accepted the situation . He has n't played any first-team games for [13 13 Villa 13] this season and he 's not the type of player I would have brought on as a substitute , " [14 14 M cCarthy 14] said . " But he surprised me in training over the last two days because of his involvement . He 's certainly is still very much part of my plans for the future . " At 24 , 25 or 26 you could get away with it , not having played first-team games . But at 36 it would be asking too much of [1 Paul 1] , " he said . Also omitted from the 20-man squad which will travel to [15 15 Vaduz 15] for Saturday 's group eight match are central defenders [16 16 Alan Kernaghan 16] and [17 17 Liam Daish 17] . [18 18 Leeds United 18] defender [19 19 Ga ry Kelly 19] is unable to travel because of a knee injury picked up in Monday 's 1-0 victory over [20 20 Wimbledon 20] at [21 21 Elland Road 21] . Since taking over from [22 22 Jack Charlton 22] in February , [23 23 McCarthy 23] has played

largely experimental sides and seen them lose five times , draw twice and win just once . Squad : [24 24 Alan Kelly 24] , [25 25 Shay Given 25] , [26 26 Denis Irwin 26] , [27 27 Phil Babb 27] , Jeff Kenna , [28 28 Curtis Fleming 28] , [29 29 Gary Breen 29] , [30 30 Ian Harte 30] , [31 31 Kenny Cunningham 31] , [32 32 Steve Staunton 32] , [33 33 Andy Townsend 33] , [34 34 Ray Houghton 34] , [35 35 Gareth Farrelly 35] , [36 36 Alan McLoughlin 36] , [37 37 Jason McAteer 37] , [38 38 Alan Moore 38] , Keith O'Neill , [39 39 Tony Cascarino 39] , [40 40 Niall Quinn 40] , [41 41 David Kelly 41] . -- Dublin Newsroom +6613377

- 1: 1438647 Paul_McGrath_(footballer)
- 2: 147575 Ireland
- 3: 126460 1998_FIFA_World_Cup
- 4: 8504 Dublin
- 5: 147575 Ireland
- 6: 1438647 Paul_McGrath_(footballer)
- 7: 195258 Mick_McCarthy
- 8: 652703 Liechtenstein_national_football_team
- 9: 42173 Aston_Villa_F.C.
- 10: 14560 Republic_of_Ireland
- 11: 5321 Czech_Republic
- 12: 23844 Prague
- 13: 42173 Aston_Villa_F.C.
- 14: 195258 Mick_McCarthy
- 15: 59036 Vaduz
- 16: 2621992 Alan_Kernaghan
- 17: 2189984 Liam_Daish
- 18: 7609513 Leeds_United_A.F.C.
- 19: 2767206 Gary_Kelly_(footballer_born_1974)
- 20: 740757 Wimbledon_F.C.
- 21: 706736 Elland_Road
- 22: 180733 Jack_Charlton
- 23: 195258 Mick_McCarthy
- 24: 8298264 Alan_Kelly,_Sr.
- 25: 610477 Shay_Given
- 26: 1114589 Denis_Irwin
- 27: 4248883 Phil_Babb
- 28: 5278717 Curtis_Fleming
- 29: 1234641 Gary_Breen
- 30: 3489770 Ian_Harte
- 31: 1331171 Kenny_Cunningham
- 32: 1227638 Steve_Staunton
- 33: 2255163 Andy_Townsend
- 34: 1219053 Ray_Houghton
- 35: 3501484 Gareth_Farrelly
- 36: 3163529 Alan_McLoughlin
- 37: 1929564 Jason_McAteer
- 38: 7211100 Alan_Moore_(footballer)
- 39: 2429090 Tony_Cascarino
- 40: 605913 Niall_Quinn
- 41: 3589053 David_Kelly_(association_footballer)

```
1: 1438647 Paul_McGrath_(footballer)
   [pem_pos=0, final_scores=0.81378895,
final_local_scores=0.8331951, pem=0.0, similarity=28.953917]
2: 147575 Ireland
   [pem_pos=0, final_scores=0.17413042,
final_local_scores=0.20527396, pem=-0.5653214,
similarity=13.289803]
3: 126460 1998_FIFA_World_Cup
   [pem_pos=18, final_scores=0.21580912,
final_local_scores=0.19944143, pem=-4.2172847,
similarity=32.975693]
4: 8504 Dublin
   [pem_pos=0, final_scores=0.45232207,
final_local_scores=0.49184048, pem=-0.23150389,
similarity=20.026169]
5: 147575 Ireland
   [pem_pos=0, final_scores=0.20497803,
final_local_scores=0.24418467, pem=-0.5653214,
similarity=14.450976]
6: 1438647 Paul_McGrath_(footballer)
   [pem_pos=0, final_scores=0.5628352,
final_local_scores=0.55551565, pem=0.0, similarity=20.667408]
7: 195258 Mick_McCarthy
   [pem_pos=0, final_scores=0.55900675,
final_local_scores=0.5312543, pem=0.0, similarity=19.943401]
8: 652703 Liechtenstein_national_football_team
   [pem_pos=1, final_scores=0.47975177,
final_local_scores=0.50662565, pem=-1.9919682,
similarity=30.041065]
9: 42173 Aston_Villa_F.C.
   [pem_pos=0, final_scores=0.5524266,
final_local_scores=0.51307446, pem=0.0, similarity=19.400879]
10: 14560 Republic_of_Ireland
    [pem_pos=3, final_scores=0.59317166,
final_local_scores=0.66763085, pem=-2.2576945,
similarity=36.29084]
11: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.39425427,
final_local_scores=0.48452103, pem=-0.27794555,
similarity=20.060299]
12: 23844 Prague
    [pem_pos=0, final_scores=0.47191447,
final_local_scores=0.59466195, pem=-0.05392049,
similarity=22.12884]
13: 42173 Aston_Villa_F.C.
    [pem_pos=7, final_scores=0.38008696,
final_local_scores=0.32520276, pem=-3.199224,
similarity=31.19229]
14: 195258 Mick_McCarthy
    [pem_pos=0, final_scores=1.2272556,
final_local_scores=1.263807, pem=0.0, similarity=41.804237]
15: 59036 Vaduz
```

[pem_pos=0, final_scores=0.27170563,
final_local_scores=0.3764838, pem=0.0, similarity=15.32474]
16: 2621992 Alan_Kernaghan
[pem_pos=0, final_scores=0.6294193,
final_local_scores=0.66484153, pem=0.0, similarity=23.92991]
17: 2189984 Liam_Daish
[pem_pos=0, final_scores=0.5411268,
final_local_scores=0.57253975, pem=0.0, similarity=21.175442]
18: 7609513 Leeds_United_A.F.C.
[pem_pos=0, final_scores=0.49459624,
final_local_scores=0.4500926, pem=0.0, similarity=17.521374]
19: 2767206 Gary_Kelly_(footballer_born_1974)
[pem_pos=0, final_scores=0.3685484,
final_local_scores=0.38084674, pem=0.0, similarity=15.454939]
20: 740757 Wimbledon_F.C.
[pem_pos=1, final_scores=0.66752,
final_local_scores=0.6637681, pem=-1.8899087,
similarity=34.175495]
21: 706736 Elland_Road
[pem_pos=0, final_scores=0.4928478,
final_local_scores=0.46972895, pem=0.0, similarity=18.10736]
22: 180733 Jack_Charlton
[pem_pos=0, final_scores=0.5497153,
final_local_scores=0.5217631, pem=0.0, similarity=19.660164]
23: 195258 Mick_McCarthy
[pem_pos=0, final_scores=1.1983882,
final_local_scores=1.2323208, pem=0.0, similarity=40.86462]
24: 8298264 Alan_Kelly,_Sr.
[pem_pos=2, final_scores=0.37368912,
final_local_scores=0.3782336, pem=-1.3307742,
similarity=22.613916]
25: 610477 Shay_Given
[pem_pos=0, final_scores=0.7463785,
final_local_scores=0.7527865, pem=0.0, similarity=26.554363]
26: 1114589 Denis_Irwin
[pem_pos=0, final_scores=0.63558644,
final_local_scores=0.63396657, pem=0.0, similarity=23.00854]
27: 4248883 Phil_Babb
[pem_pos=0, final_scores=0.50520974,
final_local_scores=0.5124412, pem=0.0, similarity=19.38198]
28: 5278717 Curtis_Fleming
[pem_pos=0, final_scores=0.5077326,
final_local_scores=0.52915895, pem=0.0, similarity=19.88087]
29: 1234641 Gary_Breen
[pem_pos=0, final_scores=0.6033531,
final_local_scores=0.6107926, pem=0.0, similarity=22.316982]
30: 3489770 Ian_Harte
[pem_pos=0, final_scores=0.5855058,
final_local_scores=0.59947616, pem=0.0, similarity=21.979277]
31: 1331171 Kenny_Cunningham
[pem_pos=0, final_scores=0.61524075,
final_local_scores=0.6282076, pem=0.0, similarity=22.836681]

```
32: 1227638 Steve_Staunton
    [pem_pos=0, final_scores=0.7600382,
final_local_scores=0.75707173, pem=0.0, similarity=26.682243]
33: 2255163 Andy_Townsend
    [pem_pos=0, final_scores=0.50202435,
final_local_scores=0.5016615, pem=0.0, similarity=19.060291]
34: 1219053 Ray_Houghton
    [pem_pos=0, final_scores=0.64204246,
final_local_scores=0.64562845, pem=0.0, similarity=23.356554]
35: 3501484 Gareth_Farrelly
    [pem_pos=0, final_scores=0.5790798,
final_local_scores=0.6164933, pem=0.0, similarity=22.487103]
36: 3163529 Alan_McLoughlin
    [pem_pos=0, final_scores=0.5742303,
final_local_scores=0.5934189, pem=0.0, similarity=21.798517]
37: 1929564 Jason_McAteer
    [pem_pos=0, final_scores=0.5819992,
final_local_scores=0.5786559, pem=0.0, similarity=21.35796]
38: 7211100 Alan_Moore_(footballer)
    [pem_pos=1, final_scores=0.6259545,
final_local_scores=0.6810255, pem=-1.3755322,
similarity=31.89323]
39: 2429090 Tony_Cascarino
    [pem_pos=0, final_scores=0.6893696,
final_local_scores=0.7215815, pem=0.0, similarity=25.623144]
40: 605913 Niall_Quinn
    [pem_pos=0, final_scores=0.6054592,
final_local_scores=0.5991997, pem=0.0, similarity=21.971027]
41: 3589053 David_Kelly_(association_footballer)
    [pem_pos=2, final_scores=0.4430297,
final_local_scores=0.45030266, pem=-1.4605583,
similarity=25.470387]
1: 1438647 Paul_McGrath_(footballer)
    [pem_pos=0, final_scores=0.056650296,
final_local_scores=0.0025855005, pem=0.0,
similarity=4.1668696]
```

Sample: b'553_S.'

S. African apartheid killer convicted of six murders . [1 1 P
RETORIA 1] 1996-08-27 [2 2 South African 2] apartheid killer [3 3 Eugene de Kock 3] was found guilty of murder and attempted murder on Tuesday , a day after he was convicted of five other murders . De Kock , 48 , a former police colonel who commanded a hit-squad that wiped out opponents of apartheid , is the most senior servant of white rule yet to face justice .

```
1: 25002 Pretoria
2: 17416221 South_Africa
3: 8324621 Eugene_de_Kock
1: 25002 Pretoria
    [pem_pos=0, final_scores=0.4238817,
final_local_scores=0.4191026, pem=0.0, similarity=16.596571]
```

```
2: 17416221 South_Africa
    [pem_pos=0, final_scores=0.4857183,
final_local_scores=0.51069206, pem=-0.1389451,
similarity=20.085386]
3: 8324621 Eugene_de_Kock
    [pem_pos=0, final_scores=0.43247753,
final_local_scores=0.5372103, pem=0.0, similarity=20.12114]
```

Sample: b'554_Sudanese'

[1 1 Sudanese 1] rebels say missionaries should be freed . Peter Smerdon [2 2 NAIROBI 2] 1996-08-27 The main rebel group in south [3 3 Sudan 3] said on Tuesday it was trying to arrange the release of six [4 4 Roman Catholic 4] missionaries , including three [5 5 Australian 5] nuns , held for nearly two weeks . George Garang , [6 6 Nairobi 6] spokesman for the Sudan People 's Liberation Army (SPLA) , said it was urgently trying to contact SPLA commander Nuour Marial at Mapourdit in the south to free the six . " The movement is making arrangements for them to be set free . This is a decision of the leadership , " Garang said . " Commander Nuour Marial is a soldier so he must accept the leadership 's decision . But communications at this time of year are very difficult because of rains and a lack of power , " he added . The Catholic Information Office in [7 7 N airobi 7] said on Monday that four of the six had been charged by the SPLA with spying , spreading [8 8 Islam 8] and hindering recruitment into the rebel group . " These charges are the interpretation of the church , " Garang said on Tuesday . " We have no idea why they are being held . We are still trying to establish contact with the local commander . " Asked whether this meant the commander was out of control , Garang said the rebel movement was working on the problem . He said he believed all six were being held in the mission compound at Mapourdit and were reported to be in good health . The Catholic Information office in [9 9 Nairobi 9] said on Monday that [10 10 Australian 10] Sisters Moira Lynch , 73 , and Mary Batchelor , 68 , [11 11 American 11] Father Michael Barton , 48 , and [12 12 Sudanese 12] Father Raphael Riel , 48 , were held in a prison in south [13 13 Sud an 13] by the SPLA . It said [14 14 Australian 14] Sister Maureen Carey , 52 , and [15 15 Italian 15] Brother Raniero Iacomella , 28 , were held inside the compound . The church in [16 16 Australia 16] said on Monday Lynch , Batchelor , Barton and Riel were held in a prison until the weekend , when they were moved to join the other captives at the compound . The Catholic Information Office said the SPLA in the [17 17 Kenyan 17] capital had attributed the detentions of the six to a local commander and had promised they would be freed by August 23 . But the church learned in a recent meeting with the local commander that no instructions to release the prisoners were received and they would be held until investigations were completed . It said last Friday

they were visited by Monsignor Caesar Mazzolari , apostolic administrator of the diocese of [18 18 Rumbek 18] in southern [19 19 Sudan 19] , and an SPLA administrator and appeared in good condition . " On August 17 the mission was surrounded (by the SPLA) and sealed off . The evening of the same day the missionaries were put in prison or isolation . Later the mission was looted , " it added . An [20 20 Australian 20] foreign ministry official said the charges against them were " fairly bizarre " and a matter for concern . He said [21 21 A ustralian 21] diplomats in [22 22 Nairobi 22] were working with the [23 23 Roman Catholic 23] church in southern [24 24 S udan 24] and with [25 25 U.S. 25] and [26 26 Italian 26] diplomats in the region to help free the missionaries . The SPLA has fought [27 27 Khartoum 27] 's government forces in the south since 1983 for greater autonomy or independence of the mainly 1 Christian and animist region from the [28 28 Mos lem 28] , Arabised north .

1: 27421 Sudan
2: 21482 Nairobi
3: 27421 Sudan
4: 606848 Catholic_Church
5: 4689264 Australia
6: 21482 Nairobi
7: 21482 Nairobi
8: 6037917 Islam
9: 21482 Nairobi
10: 4689264 Australia
11: 3434750 United_States
12: 27421 Sudan
13: 27421 Sudan
14: 4689264 Australia
15: 14532 Italy
16: 4689264 Australia
17: 188171 Kenya
18: 1580059 Rumbek
19: 27421 Sudan
20: 4689264 Australia
21: 4689264 Australia
22: 21482 Nairobi
23: 606848 Catholic_Church
24: 27421 Sudan
25: 3434750 United_States
26: 14532 Italy
27: 17320 Khartoum
28: 6037917 Islam
1: 27968837 Christian_theology
[winner: 18337522 Christian, pem_pos=0,
final_scores=-0.17069711, final_local_scores=-0.2164731,
pem=-0.5699843, similarity=0.72938746]
RECALL MISS: 27968837 Christian_theology]
1: 27421 Sudan
[pem_pos=0, final_scores=0.56653726,

```
final_local_scores=0.5292349, pem=-0.19893567,
similarity=20.964981]
2: 21482 Nairobi
    [pem_pos=0, final_scores=0.449598,
final_local_scores=0.42287856, pem=0.0, similarity=16.709253]
3: 27421 Sudan
    [pem_pos=0, final_scores=0.5207545,
final_local_scores=0.4792242, pem=-0.2848764,
similarity=19.939922]
4: 606848 Catholic_Church
    [pem_pos=0, final_scores=0.35367745,
final_local_scores=0.3874439, pem=0.0, similarity=15.651812]
5: 4689264 Australia
    [pem_pos=0, final_scores=0.6089806,
final_local_scores=0.5939095, pem=-0.491009,
similarity=24.48334]
6: 21482 Nairobi
    [pem_pos=0, final_scores=0.49738544,
final_local_scores=0.47507906, pem=0.0, similarity=18.26702]
7: 21482 Nairobi
    [pem_pos=0, final_scores=0.4419586,
final_local_scores=0.41453367, pem=0.0, similarity=16.460224]
8: 6037917 Islam
    [pem_pos=0, final_scores=0.43884647,
final_local_scores=0.48038888, pem=-0.12106096,
similarity=19.083822]
9: 21482 Nairobi
    [pem_pos=0, final_scores=0.46303362,
final_local_scores=0.4375549, pem=0.0, similarity=17.147224]
10: 4689264 Australia
    [pem_pos=0, final_scores=0.63766617,
final_local_scores=0.6252442, pem=-0.491009,
similarity=25.41843]
11: 3434750 United_States
    [pem_pos=0, final_scores=0.47146887,
final_local_scores=0.44753265, pem=-0.12994486,
similarity=18.151642]
12: 27421 Sudan
    [pem_pos=0, final_scores=0.47599238,
final_local_scores=0.43032837, pem=-0.19893567,
similarity=18.013412]
13: 27421 Sudan
    [pem_pos=0, final_scores=0.54348737,
final_local_scores=0.5040564, pem=-0.2848764,
similarity=20.680964]
14: 4689264 Australia
    [pem_pos=0, final_scores=0.72380275,
final_local_scores=0.7193353, pem=-0.491009,
similarity=28.226297]
15: 14532 Italy
    [pem_pos=0, final_scores=0.6838202,
final_local_scores=0.7093793, pem=-0.75948656,
```

```
similarity=29.389212]
16: 4689264 Australia
    [pem_pos=0, final_scores=0.46636456,
final_local_scores=0.43812323, pem=-0.17168477,
similarity=18.097834]
17: 188171 Kenya
    [pem_pos=0, final_scores=0.5767334,
final_local_scores=0.5593612, pem=0.0, similarity=20.782167]
18: 1580059 Rumbek
    [pem_pos=0, final_scores=0.3585806,
final_local_scores=0.43058538, pem=0.0, similarity=16.93924]
19: 27421 Sudan
    [pem_pos=0, final_scores=0.52029675,
final_local_scores=0.47872418, pem=-0.2848764,
similarity=19.925]
20: 4689264 Australia
    [pem_pos=0, final_scores=0.6205235,
final_local_scores=0.60651845, pem=-0.491009,
similarity=24.859615]
21: 4689264 Australia
    [pem_pos=0, final_scores=0.6428501,
final_local_scores=0.6309068, pem=-0.491009,
similarity=25.587414]
22: 21482 Nairobi
    [pem_pos=0, final_scores=0.48971123,
final_local_scores=0.46669608, pem=0.0, similarity=18.016855]
23: 606848 Catholic_Church
    [pem_pos=0, final_scores=0.4192577,
final_local_scores=0.45908034, pem=0.0, similarity=17.789585]
24: 27421 Sudan
    [pem_pos=0, final_scores=0.5492752,
final_local_scores=0.5103787, pem=-0.2848764,
similarity=20.869635]
25: 3434750 United_States
    [pem_pos=0, final_scores=0.500083,
final_local_scores=0.47878927, pem=-0.14075686,
similarity=19.143196]
26: 14532 Italy
    [pem_pos=1, final_scores=0.68669593,
final_local_scores=0.7125206, pem=-0.75948656,
similarity=29.482954]
27: 17320 Khartoum
    [pem_pos=0, final_scores=0.32097435,
final_local_scores=0.29751322, pem=-0.03966028,
similarity=13.18378]
28: 6037917 Islam
    [pem_pos=1, final_scores=0.3925513,
final_local_scores=0.4298184, pem=-1.2150909,
similarity=23.524204]
```

Sample: b'555_OSCE'

OSCE postpones Bosnian municipal elections . [1 1 SARAJEVO 1]

1996-08-27 The [2 2 U.S. 2] diplomat in charge of elections in Bosnia announced on Tuesday that voting for municipal assemblies would be postponed because of irregularities by the [3 3 Serbs 3] in registering voters . Ambassador Robert Frowick , representing the Organisation for Security and Cooperation in Europe (OSCE) , told reporters that municipal polls due on September 14 with other Bosnian elections would be put off . " I have made a chairman 's decision that it is not feasible to hold municipal elections on September 14 , " said Frowick . He said no exact date had been set but it was possible the local elections would take place in the spring of 1997 . According to OSCE officials , [4 4 Serb 4] authorities have pressed their refugees to register to vote in towns now under [5 5 Serb 5] control , but which used to have [6 6 Moslem 6] majorities . Human rights workers say authorities in [7 7 Serbia 7] and [8 8 Bosnian Serb 8] territory have conducted a well-organised campaign to coerce refugees into registering only on [9 9 Serb 9] territory and failed to inform them of their rights under the 1 Dayton peace agreement . Diplomats say the effect of the electoral engineering would be to establish political control over districts they conquered and ethnically cleansed in war . The response of the Bosnian Serbs to the OSCE 's announcement was not immediately clear . But [10 10 Bosnian Serb 10] leaders have hinted they would boycott the poll if the municipal elections were postponed , or go ahead with their own . The [1 11 Bosnian Serb 11] cabinet , in a letter to the OSCE , said on Monday that any delay of local elections would be " a direct and flagrant violation fo the 2 Dayton agreement " . The [12 12 Serbs 12] , who administer half of Bosnia in a [13 13 Serb 13] republic , said they had met all conditions for holding the September elections . Diplomats fear that the crisis could cast doubt over the entire election process , which already appears set to confirm Bosnia 's ethnic partition rather than its reintegration as the 3 Dayton peace agreement had planned .

1: 26786 Sarajevo
2: 3434750 United_States
3: 29265 Serbia
4: 29265 Serbia
5: 29265 Serbia
6: 6037917 Islam
7: 29265 Serbia
8: 347422 Republika_Srpska
9: 29265 Serbia
10: 347422 Republika_Srpska
11: 347422 Republika_Srpska
12: 29265 Serbia
13: 29265 Serbia
1: 168269 Dayton_Agreement
[winner: 8253 Dayton,_Ohio, pem_pos=0,
final_scores=-0.18866985, final_local_scores=-0.18088204,

```
pem=-0.9297534, similarity=3.7479756 |  
    RECALL MISS: 168269 Dayton_Agreement]  
2: 168269 Dayton_Agreement  
    [winner: 8253 Dayton,_Ohio, pem_pos=0,  
final_scores=-0.118838295, final_local_scores=-0.10460168,  
pem=-0.9297534, similarity=6.0243335 |  
    RECALL MISS: 168269 Dayton_Agreement]  
3: 168269 Dayton_Agreement  
    [winner: 8253 Dayton,_Ohio, pem_pos=0,  
final_scores=-0.034656048, final_local_scores=0.009356007,  
pem=-0.9297534, similarity=9.425058 |  
    RECALL MISS: 168269 Dayton_Agreement]  
1: 26786 Sarajevo  
    [pem_pos=0, final_scores=0.4122302,  
final_local_scores=0.38301343, pem=-0.0017846047,  
similarity=15.529303]  
2: 3434750 United_States  
    [pem_pos=0, final_scores=0.41918594,  
final_local_scores=0.44400233, pem=-0.14075686,  
similarity=18.105087]  
3: 29265 Serbia  
    [pem_pos=4, final_scores=0.5649236,  
final_local_scores=0.5019862, pem=-2.3277044,  
similarity=31.728401]  
4: 29265 Serbia  
    [pem_pos=1, final_scores=0.604521,  
final_local_scores=0.5452404, pem=-2.1726613,  
similarity=32.176044]  
5: 29265 Serbia  
    [pem_pos=1, final_scores=0.64336014,  
final_local_scores=0.5876663, pem=-2.1726613,  
similarity=33.442116]  
6: 6037917 Islam  
    [pem_pos=1, final_scores=0.40991753,  
final_local_scores=0.51303804, pem=-1.2150909,  
similarity=26.007645]  
7: 29265 Serbia  
    [pem_pos=0, final_scores=0.59829813,  
final_local_scores=0.53844285, pem=-0.34077588,  
similarity=22.011114]  
8: 347422 Republika_Srpska  
    [pem_pos=0, final_scores=0.43779427,  
final_local_scores=0.4227664, pem=0.0, similarity=16.705906]  
9: 29265 Serbia  
    [pem_pos=1, final_scores=0.5955896,  
final_local_scores=0.5354842, pem=-2.1726613,  
similarity=31.8849]  
10: 347422 Republika_Srpska  
    [pem_pos=0, final_scores=0.3869241,  
final_local_scores=0.3671984, pem=0.0, similarity=15.0476465]  
11: 347422 Republika_Srpska  
    [pem_pos=0, final_scores=0.42371994,
```

```
final_local_scores=0.40739226, pem=0.0, similarity=16.24711]
12: 29265 Serbia
    [pem_pos=4, final_scores=0.52246505,
final_local_scores=0.45560676, pem=-2.3277044,
similarity=30.344345]
13: 29265 Serbia
    [pem_pos=1, final_scores=0.63307726,
final_local_scores=0.5764338, pem=-2.1726613,
similarity=33.106915]
```

Sample: b'556_New'

New talks in [1 1 Chechnya 1] as [2 2 Lebed 2] waits for [3 3 Yeltsin 3] . Dmitry Kuznets [4 4 NOVYE ATAGI 4] , [5 5 Russia 5] 1996-08-27 [6 6 Russian 6] and rebel military commanders finally met in [7 7 Chechnya 7] on Tuesday for delayed talks aimed at finalising a ceasefire arranged last week by President [8 8 Boris Yeltsin 8] 's envoy [9 9 Alexander Lebed 9] . The [10 10 Russian 10] army commander in the region , General Vyacheslav Tikhomirov , arrived at the rebel-held village of [11 11 Novye Atagi 11] , some 20 km (12 miles) south of the [12 12 Chechen 12] capital [13 13 Grozny 13] for discussions with rebel chief-of-staff [14 14 Aslan Maskhadov 1 4] . But [15 15 Lebed 15] himself , the [16 16 Kremlin 16] security chief , is still waiting back in [17 17 Moscow 17] to meet [18 18 Yeltsin 18] over his plans for a lasting political settlement in [19 19 Chechnya 19] . [20 20 Itar-Tass 20] news agency quoted the [21 21 Kremlin 21] press service as saying [22 22 Yeltsin 22] , who left for a state holiday home near [23 23 Moscow 23] on Monday , would hold no working meetings on Tuesday . [24 24 Lebed 24] interrupted talks with [25 25 Chechnya 25] 's separatists on a political deal on Sunday , saying he had to consult with [26 26 Yeltsin 26] . After a meeting failed to materialise on Monday , [27 27 Lebed 27] 's spokesman said he might meet the president on Tuesday . But [28 28 Yeltsin 28] 's spokesman rebuffed the suggestion , saying the president had left [29 29 Moscow 29] for a holiday near the capital . The [30 30 Russians 30] postponed the talks after a [31 31 Chechen 31] band disarmed a column of [32 32 interior ministry 32] troops on Sunday . The [33 33 Chechens 33] said a renegade group seized the weapons and said on Monday they had all been returned . The [3 4 34 Russian 34] command insisted that not all the weapons were the same as those taken . [35 35 Tass 35] said the weapons and the practical implementation of the ceasefire signed by [36 36 Lebed 36] and [37 37 Maskhadov 37] last Thursday would be on the agenda of today 's talks . Neither spoke to reporters before the meeting , which started around 10.45 a.m. (0645 [38 38 GMT 38]) . Also in [39 39 Novye Atagi 39] on Tuesday morning , was Tim Guldmann , the [40 40 Swiss 40] diplomat who heads the [41 41 Chechnya 41] mission of the Organisation for Security and Cooperation in Europe (OSCE) . Guldmann , who helped broker an earlier truce in

May , was not taking part in the Tikhomirov-Maskhadov talks . [42 42 Lebed 42] 's peace mission this month has stopped some of the worst fighting of the 20-month-old conflict . However , tension on the ground indicates that it could falter if the momentum for a settlement is not maintained . Three [43 43 Russian 43] servicemen were wounded in a total of six shooting incidents overnight , [44 44 Itar-Tass 44] news agency quoted the [45 45 Russian 45] military as saying on Tuesday morning . RIA news agency quoted an army source accusing rebel fighters of failing to turn up for joint Russian-Chechen police patrols in some districts of the capital [46 46 Grozny 46] on Tuesday . But the separatist command told [47 47 Interfax 47] news agency the patrols , part of the truce brokered by [48 48 Lebed 48] last week , would begin on Tuesday after delays for " technical reasons " . [49 49 Yeltsin 49] 's spokesman said he might meet officials during his break , but indicated [50 50 Lebed 50] was low on the list by saying [51 51 Yeltsin 51] would need time to study the proposals before talking to him . [52 52 Russian 52] news agencies also quoted the [53 53 Kremlin 53] spokesman as saying that [54 54 Lebed 54] 's representatives had not sought a meeting , hinting at an attempt by the president to put his popular and outspoken protege in his place with a lesson on protocol . [55 55 Yeltsin 55] , 65 , has kept a low profile since he was reelected in July , prompting new speculation that the two heart attacks he suffered last year and a rumoured drinking problem could be taking their toll , weakening his grip on affairs of state . Aides have dismissed such speculation , insisting that he simply needs a rest after his energetic election campaign . Some analysts say the [56 56 Kremlin 56] leader , whose order sending troops and tanks into [57 57 Chechnya 57] in 1994 started [58 58 Russia 58] 's ill-fated military campaign , could merely be reluctant to put his name to a peace process which might fall apart . But [59 59 Lebed 59] , who has no real power without his boss and has hinted at dark forces in [60 60 Moscow 60] working against him , appears to think a deal will not stick without strong backing from [61 61 Yeltsin 61] . His proposals have not been spelled out but are expected to involve a compromise between the separatists ' demand for independence and [62 62 Moscow 62] 's insistence that [63 63 Chechnya 63] remain part of the [64 64 Russian Federation 64] .

- 1: 6095 Chechnya
- 2: 431579 Alexander_Lebed
- 3: 23278520 Boris_Yeltsin
- 4: 4813770 Novye_Atagi
- 5: 25391 Russia
- 6: 25391 Russia
- 7: 6095 Chechnya
- 8: 23278520 Boris_Yeltsin
- 9: 431579 Alexander_Lebed

10: 25391 Russia
11: 4813770 Novye_Atagi
12: 6095 Chechnya
13: 140790 Grozny
14: 340963 Aslan_Maskhadov
15: 431579 Alexander_Lebed
16: 9787373 Government_of_Russia
17: 19004 Moscow
18: 23278520 Boris_Yeltsin
19: 6095 Chechnya
20: 434852 Information_Telegraph_Agency_of_Russia
21: 9787373 Government_of_Russia
22: 23278520 Boris_Yeltsin
23: 19004 Moscow
24: 431579 Alexander_Lebed
25: 6095 Chechnya
26: 23278520 Boris_Yeltsin
27: 431579 Alexander_Lebed
28: 23278520 Boris_Yeltsin
29: 19004 Moscow
30: 25391 Russia
31: 6095 Chechnya
32: 354620 Interior_ministry
33: 6095 Chechnya
34: 25391 Russia
35: 2711386 Telegraph_Agency_of_the_Soviet_Union
36: 431579 Alexander_Lebed
37: 340963 Aslan_Maskhadov
38: 12701 Greenwich_Mean_Time
39: 4813770 Novye_Atagi
40: 26748 Switzerland
41: 6095 Chechnya
42: 431579 Alexander_Lebed
43: 25391 Russia
44: 434852 Information_Telegraph_Agency_of_Russia
45: 25391 Russia
46: 140790 Grozny
47: 4921404 Interfax
48: 431579 Alexander_Lebed
49: 23278520 Boris_Yeltsin
50: 431579 Alexander_Lebed
51: 23278520 Boris_Yeltsin
52: 25391 Russia
53: 9787373 Government_of_Russia
54: 431579 Alexander_Lebed
55: 23278520 Boris_Yeltsin
56: 381075 Moscow_Kremlin
57: 6095 Chechnya
58: 25391 Russia
59: 431579 Alexander_Lebed
60: 19004 Moscow
61: 23278520 Boris_Yeltsin

62: 19004 Moscow
63: 6095 Chechnya
64: 25391 Russia
1: 6095 Chechnya
[pem_pos=0, final_scores=0.5745694,
final_local_scores=0.5541622, pem=0.0, similarity=20.627018]
2: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.6353994,
final_local_scores=0.6184583, pem=0.0, similarity=22.545742]
3: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.53492993,
final_local_scores=0.5044164, pem=0.0, similarity=19.142504]
4: 4813770 Novye_Atagi
[pem_pos=0, final_scores=0.37322706,
final_local_scores=0.46061713, pem=0.0, similarity=17.835447]
5: 25391 Russia
[pem_pos=0, final_scores=0.8054158,
final_local_scores=0.8365323, pem=-0.20299491,
similarity=30.157423]
6: 25391 Russia
[pem_pos=1, final_scores=0.7562588,
final_local_scores=0.7828357, pem=-1.0515022,
similarity=33.169323]
7: 6095 Chechnya
[pem_pos=0, final_scores=0.58337665,
final_local_scores=0.56378275, pem=0.0, similarity=20.914114]
8: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.67433804,
final_local_scores=0.6566986, pem=0.0, similarity=23.686909]
9: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.73012054,
final_local_scores=0.72030246, pem=0.0, similarity=25.584974]
10: 25391 Russia
[pem_pos=1, final_scores=0.898174,
final_local_scores=0.9378565, pem=-1.0515022,
similarity=37.795456]
11: 4813770 Novye_Atagi
[pem_pos=0, final_scores=0.5400834,
final_local_scores=0.6428823, pem=0.0, similarity=23.274603]
12: 6095 Chechnya
[pem_pos=2, final_scores=0.48112315,
final_local_scores=0.45208633, pem=-0.8617233,
similarity=22.267057]
13: 140790 Grozny
[pem_pos=0, final_scores=0.49604303,
final_local_scores=0.4866547, pem=0.0, similarity=18.61246]
14: 340963 Aslan_Maskhadov
[pem_pos=0, final_scores=0.6476359,
final_local_scores=0.6580931, pem=0.0, similarity=23.728523]
15: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.650589,
final_local_scores=0.6350506, pem=0.0, similarity=23.04089]

16: 9787373 Government_of_Russia
[pem_pos=3, final_scores=0.55006486,
final_local_scores=0.5488955, pem=-2.0052693,
similarity=31.374815]
17: 19004 Moscow
[pem_pos=0, final_scores=0.63568324,
final_local_scores=0.6410987, pem=-0.106582455,
similarity=23.80099]
18: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.6253521,
final_local_scores=0.6031888, pem=0.0, similarity=22.09007]
19: 6095 Chechnya
[pem_pos=0, final_scores=0.59352726,
final_local_scores=0.57487077, pem=0.0, similarity=21.245003]
20: 434852 Information_Telegraph_Agency_of_Russia
[pem_pos=0, final_scores=0.7001917,
final_local_scores=0.75452065, pem=0.0, similarity=26.606113]
21: 9787373 Government_of_Russia
[pem_pos=3, final_scores=0.45558745,
final_local_scores=0.4456933, pem=-2.0052693,
similarity=28.295057]
22: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.6055451,
final_local_scores=0.58155274, pem=0.0, similarity=21.444407]
23: 19004 Moscow
[pem_pos=0, final_scores=0.6101571,
final_local_scores=0.6132152, pem=-0.106582455,
similarity=22.968891]
24: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.6869033,
final_local_scores=0.67471844, pem=0.0, similarity=24.224657]
25: 6095 Chechnya
[pem_pos=0, final_scores=0.63264585,
final_local_scores=0.6176019, pem=0.0, similarity=22.520185]
26: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.58273524,
final_local_scores=0.5566365, pem=0.0, similarity=20.700857]
27: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.65128314,
final_local_scores=0.6358089, pem=0.0, similarity=23.063519]
28: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.6644098,
final_local_scores=0.6458535, pem=0.0, similarity=23.36327]
29: 19004 Moscow
[pem_pos=0, final_scores=0.5507906,
final_local_scores=0.5483663, pem=-0.106582455,
similarity=21.03367]
30: 25391 Russia
[pem_pos=1, final_scores=0.6790512,
final_local_scores=0.6984982, pem=-1.8802736,
similarity=35.15951]
31: 6095 Chechnya

[pem_pos=2, final_scores=0.48172718,
final_local_scores=0.45274615, pem=-0.8617233,
similarity=22.286747]
32: 354620 Interior_ministry
[pem_pos=0, final_scores=0.5036246,
final_local_scores=0.5373796, pem=-0.57122487,
similarity=23.232601]
33: 6095 Chechnya
[pem_pos=2, final_scores=0.47363818,
final_local_scores=0.44391018, pem=-0.96860826,
similarity=22.60432]
34: 25391 Russia
[pem_pos=1, final_scores=0.87826115,
final_local_scores=0.91610473, pem=-1.0515022,
similarity=37.14634]
35: 2711386 Telegraph_Agency_of_the_Soviet_Union
[pem_pos=1, final_scores=0.23625086,
final_local_scores=0.2512105, pem=-0.5751835,
similarity=14.7142725]
36: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.7262322,
final_local_scores=0.71767926, pem=0.0, similarity=25.506693]
37: 340963 Aslan_Maskhadov
[pem_pos=0, final_scores=0.5582755,
final_local_scores=0.5604804, pem=0.0, similarity=20.815567]
38: 12701 Greenwich_Mean_Time
[pem_pos=0, final_scores=0.6261973,
final_local_scores=0.7528652, pem=0.0, similarity=26.556711]
39: 4813770 Novye_Atagi
[pem_pos=0, final_scores=0.3830734,
final_local_scores=0.47137272, pem=0.0, similarity=18.156416]
40: 26748 Switzerland
[pem_pos=0, final_scores=0.6652029,
final_local_scores=0.75539035, pem=-0.1392413,
similarity=27.389284]
41: 6095 Chechnya
[pem_pos=0, final_scores=0.58250314,
final_local_scores=0.5628286, pem=0.0, similarity=20.885641]
42: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.6744248,
final_local_scores=0.66108763, pem=0.0, similarity=23.817886]
43: 25391 Russia
[pem_pos=1, final_scores=0.8288495,
final_local_scores=0.86213005, pem=-1.0515022,
similarity=35.535625]
44: 434852 Information_Telegraph_Agency_of_Russia
[pem_pos=0, final_scores=0.71554416,
final_local_scores=0.7712909, pem=0.0, similarity=27.106571]
45: 25391 Russia
[pem_pos=1, final_scores=0.8546323,
final_local_scores=0.8902938, pem=-1.0515022,
similarity=36.376087]

46: 140790 Grozny
[pem_pos=0, final_scores=0.5520479,
final_local_scores=0.5478315, pem=0.0, similarity=20.438097]
47: 4921404 Interfax
[pem_pos=0, final_scores=0.5742962,
final_local_scores=0.6110431, pem=0.0, similarity=22.324457]
48: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.65197915,
final_local_scores=0.63656914, pem=0.0, similarity=23.086206]
49: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.69863266,
final_local_scores=0.6832368, pem=0.0, similarity=24.47886]
50: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.7265005,
final_local_scores=0.7179724, pem=0.0, similarity=25.515442]
51: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.6913896,
final_local_scores=0.6753248, pem=0.0, similarity=24.242752]
52: 25391 Russia
[pem_pos=1, final_scores=0.76382846,
final_local_scores=0.79110444, pem=-1.0515022,
similarity=33.41608]
53: 9787373 Government_of_Russia
[pem_pos=3, final_scores=0.39232427,
final_local_scores=0.37658787, pem=-2.0052693,
similarity=26.232813]
54: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.6405373,
final_local_scores=0.62407064, pem=0.0, similarity=22.713226]
55: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.6453015,
final_local_scores=0.6249806, pem=0.0, similarity=22.740381]
56: 381075 Moscow_Kremlin
[pem_pos=1, final_scores=0.32150275,
final_local_scores=0.33370578, pem=-0.76193243,
similarity=18.191666]
57: 6095 Chechnya
[pem_pos=0, final_scores=0.514194,
final_local_scores=0.48821127, pem=0.0, similarity=18.65891]
58: 25391 Russia
[pem_pos=0, final_scores=0.84230584,
final_local_scores=0.87682897, pem=-0.20299491,
similarity=31.359955]
59: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.73423386,
final_local_scores=0.7264199, pem=0.0, similarity=25.767532]
60: 19004 Moscow
[pem_pos=0, final_scores=0.5055148,
final_local_scores=0.4989094, pem=-0.106582455,
similarity=19.557777]
61: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.6271822,

```
final_local_scores=0.605188, pem=0.0, similarity=22.14973]
62: 19004 Moscow
    [pem_pos=0, final_scores=0.41379112,
final_local_scores=0.3987152, pem=-0.106582455,
similarity=16.567781]
63: 6095 Chechnya
    [pem_pos=0, final_scores=0.57507163,
final_local_scores=0.5547108, pem=0.0, similarity=20.64339]
64: 25391 Russia
    [pem_pos=0, final_scores=0.4822833,
final_local_scores=0.48355907, pem=0.0, similarity=18.52008]
```

Sample: b'557_Slovenia'

[1 1 Slovenia 1] and [2 2 Poland 2] target EU , [3 3 NATO 3] membership . [4 4 LJUBLJANA 4] 1996-08-27 [5 5 Slovenia 5] and [6 6 Poland 6] pledged to intensify cooperation on Tuesday and reinforced their determination to join the [7 7 European Union 7] and [8 8 NATO 8] at the earliest possible date . [9 9 Polish 9] President [10 10 Aleksander Kwasniewski 10] and his [11 11 Slovenian 11] counterpart , [12 12 Milan Kucan 12] , met for talks at the start of a two-day visit to [13 13 Slovenia 13] by [14 14 Kwasniewski 14] . It was their fourth meeting this year . They said in a statement they agreed to have regular telephone contact to discuss progress in strengthening ties with the West . " We expect our cooperation will help both countries towards entering the [15 15 European Union 15] and [16 16 NATO 16] , " [17 17 Kwasniewski 17] said . " We have similar ambitions as far as our internal development and international life is concerned , " Kucan said . [18 18 Poland 18] and [19 19 Slovenia 19] are hoping to be among the first group of former eastern bloc countries to join the [20 20 European Union 20] and [21 21 NATO 21] . They have already signed an association agreement with the [22 22 European Union 22] and are both part of the Central European Free Trade Area , which also comprises [23 23 Hungary 23] , [24 24 Slovakia 24] and the [25 25 Czech Republic 25] . [26 26 Slovenia 26] 's trade with [27 27 Poland 27] rose to \$ 142.3 million in 1995 from \$ 118.8 million in 1994 . During his visit to [28 28 Slovenia 28] , [29 29 Kwasniewski 29] is also scheduled to meet Prime Minister [30 30 Janez Drnovsek 30] , representatives of [31 31 Slovenian 31] political parties and representatives of the Chamber of Economy .

```
1: 27338 Slovenia
2: 22936 Poland
3: 21133 NATO
4: 42374 Ljubljana
5: 27338 Slovenia
6: 22936 Poland
7: 9317 European_Union
8: 21133 NATO
9: 22936 Poland
```

```
10: 150186 Aleksander_Kwaśniewski
11: 27338 Slovenia
12: 429893 Milan_Kučan
13: 27338 Slovenia
14: 150186 Aleksander_Kwaśniewski
15: 9317 European_Union
16: 21133 NATO
17: 150186 Aleksander_Kwaśniewski
18: 22936 Poland
19: 27338 Slovenia
20: 9317 European_Union
21: 21133 NATO
22: 9317 European_Union
23: 13275 Hungary
24: 26830 Slovakia
25: 5321 Czech_Republic
26: 27338 Slovenia
27: 22936 Poland
28: 27338 Slovenia
29: 150186 Aleksander_Kwaśniewski
30: 162779 Janez_Drnovšek
31: 27338 Slovenia
1: 27338 Slovenia
    [pem_pos=0, final_scores=0.45412338,
final_local_scores=0.41659647, pem=-0.27845925,
similarity=18.036087]
2: 22936 Poland
    [pem_pos=0, final_scores=0.63820344,
final_local_scores=0.6097857, pem=-0.15333635,
similarity=23.1208]
3: 21133 NATO
    [pem_pos=0, final_scores=0.43879664,
final_local_scores=0.44728202, pem=0.0, similarity=17.4375]
4: 42374 Ljubljana
    [pem_pos=0, final_scores=0.42773658,
final_local_scores=0.40917605, pem=0.0, similarity=16.300343]
5: 27338 Slovenia
    [pem_pos=0, final_scores=0.45772785,
final_local_scores=0.42053378, pem=-0.27845925,
similarity=18.153584]
6: 22936 Poland
    [pem_pos=0, final_scores=0.6050779,
final_local_scores=0.5736011, pem=-0.15333635,
similarity=22.040981]
7: 9317 European_Union
    [pem_pos=0, final_scores=0.49817568,
final_local_scores=0.49661654, pem=-0.005814796,
similarity=18.941362]
8: 21133 NATO
    [pem_pos=0, final_scores=0.41370112,
final_local_scores=0.419869, pem=0.0, similarity=16.619442]
9: 22936 Poland
```

```
[pem_pos=0, final_scores=0.5264122,
final_local_scores=0.48767072, pem=-0.81518245,
similarity=23.07587]
10: 150186 Aleksander_Kwaśniewski
    [pem_pos=0, final_scores=0.73491305,
final_local_scores=0.7825943, pem=0.0, similarity=27.443888]
11: 27338 Slovenia
    [pem_pos=0, final_scores=0.31382924,
final_local_scores=0.2633465, pem=-0.5440836,
similarity=14.90731]
12: 429893 Milan_Kučan
    [pem_pos=0, final_scores=0.62818396,
final_local_scores=0.6974615, pem=0.0, similarity=24.903355]
13: 27338 Slovenia
    [pem_pos=0, final_scores=0.5115101,
final_local_scores=0.47928274, pem=-0.27845925,
similarity=19.906769]
14: 150186 Aleksander_Kwaśniewski
    [pem_pos=0, final_scores=0.56656975,
final_local_scores=0.59870476, pem=0.0, similarity=21.956257]
15: 9317 European_Union
    [pem_pos=0, final_scores=0.460667,
final_local_scores=0.455644, pem=-0.005814796,
similarity=17.71866]
16: 21133 NATO
    [pem_pos=0, final_scores=0.38460523,
final_local_scores=0.38808614, pem=0.0, similarity=15.670977]
17: 150186 Aleksander_Kwaśniewski
    [pem_pos=0, final_scores=0.5936317,
final_local_scores=0.62826586, pem=0.0, similarity=22.838419]
18: 22936 Poland
    [pem_pos=0, final_scores=0.6231714,
final_local_scores=0.59336555, pem=-0.15333635,
similarity=22.63079]
19: 27338 Slovenia
    [pem_pos=0, final_scores=0.4768861,
final_local_scores=0.44146132, pem=-0.27845925,
similarity=18.778103]
20: 9317 European_Union
    [pem_pos=0, final_scores=0.4941563,
final_local_scores=0.49222594, pem=-0.005814796,
similarity=18.810339]
21: 21133 NATO
    [pem_pos=0, final_scores=0.47695655,
final_local_scores=0.48896593, pem=0.0, similarity=18.68143]
22: 9317 European_Union
    [pem_pos=0, final_scores=0.458075,
final_local_scores=0.4528126, pem=-0.005814796,
similarity=17.634167]
23: 13275 Hungary
    [pem_pos=0, final_scores=0.5974414,
final_local_scores=0.57620627, pem=-0.32577527,
```

```
similarity=23.056475]
24: 26830 Slovakia
    [pem_pos=0, final_scores=0.5775519,
final_local_scores=0.5568954, pem=-0.36602598,
similarity=22.69909]
25: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.43709928,
final_local_scores=0.40277272, pem=-0.27794555,
similarity=17.620766]
26: 27338 Slovenia
    [pem_pos=0, final_scores=0.45934397,
final_local_scores=0.42229915, pem=-0.27845925,
similarity=18.206266]
27: 22936 Poland
    [pem_pos=0, final_scores=0.5600849,
final_local_scores=0.5244531, pem=-0.15333635,
similarity=20.574306]
28: 27338 Slovenia
    [pem_pos=0, final_scores=0.49267173,
final_local_scores=0.4587047, pem=-0.27845925,
similarity=19.29268]
29: 150186 Aleksander_Kwaśniewski
    [pem_pos=0, final_scores=0.6113824,
final_local_scores=0.64765584, pem=0.0, similarity=23.417055]
30: 162779 Janez_Drnovšek
    [pem_pos=0, final_scores=0.84642214,
final_local_scores=0.93737286, pem=0.0, similarity=32.06279]
31: 27338 Slovenia
    [pem_pos=0, final_scores=0.29961932,
final_local_scores=0.24782425, pem=-0.5440836,
similarity=14.444095]
```

Sample: b'558_Nationalists'

Nationalists want [1 1 Iliescu 1] ousted for [2 2 Hungary 2] pact . [3 3 BUCHAREST 3] 1996-08-27 Junior Nationalist members of [4 4 Romania 4] 's ruling coalition called on Tuesday for the impeachment of President [5 5 Ion Iliescu 5] for backing a friendship treaty with neighbouring [6 6 Hungary 6] . [7 7 Iliescu 7] 's [8 8 Party of Social Democracy 8] , the senior coalition partner , immediately dismissed the National Unity Party ([9 9 PUNR 9]) demand as crude electioneering . " It is a desperate move by the [10 10 PUNR 10] , which is losing its only reason for existing ahead of the electoral campaign , " said [11 11 PDSR 11] executive president [12 12 Adrian Nastase 12] . " This treaty is both necessary and good , " [13 13 Nastase 13] said , adding that the [14 14 PUNR 14] 's stance was threatening its position in the government . The treaty agreed unexpectedly two weeks ago will end years of disputes over the status of [15 15 Romania 15] 's large ethnic [16 16 Hungarian 16] minority . It will also boost both countries ' chances of admission to [17 17 NATO 17] and the [18 18 European Union 18] . " If they (the [19

[19 PUNR 19]) are so vexed , they could leave the government ... We might also help them to do it , if they go on like this , " he said . The [20 20 PUNR 20] holds four key ministries -- justice , transport , agriculture and communications . [21 21 PUNR 21] leader [22 22 Gheorghe Funar 22] said in a statement [23 23 Iliescu 23] , in power since the fall of communism in 1989 , should be impeached for treason for compromising on the issue of ethnic [24 24 Hungarian 24] minority rights in the treaty due to be signed next month . [25 25 Funar 25] 's call came on the eve of the official launch of [26 26 Iliescu 26] 's campaign for a new term at November 3 polls . His appeal to the opposition to back his attempt to oust [27 27 Iliescu 27] was unlikely to succeed , analysts said . [28 28 Iliescu 28] has invited political leaders to a meeting on Thursday to discuss the final form of the pact which both [29 29 Romanian 29] and [30 30 Hungarian 30] nationalists oppose for different reasons . Presidential officials were not available to comment on the call for [31 31 Iliescu 31] 's impeachment .

1: 189451 Ion_Iliescu
2: 13275 Hungary
3: 36877 Bucharest
4: 25445 Romania
5: 189451 Ion_Iliescu
6: 13275 Hungary
7: 189451 Ion_Iliescu
8: 407177 Social_Democratic_Party_(Romania)
9: 1969239 Romanian_National_Unity_Party
10: 1969239 Romanian_National_Unity_Party
11: 407177 Social_Democratic_Party_(Romania)
12: 395061 Adrian_Năstase
13: 395061 Adrian_Năstase
14: 1969239 Romanian_National_Unity_Party
15: 25445 Romania
16: 13275 Hungary
17: 21133 NATO
18: 9317 European_Union
19: 1969239 Romanian_National_Unity_Party
20: 1969239 Romanian_National_Unity_Party
21: 1969239 Romanian_National_Unity_Party
22: 1184425 Gheorghe_Funar
23: 189451 Ion_Iliescu
24: 13275 Hungary
25: 1184425 Gheorghe_Funar
26: 189451 Ion_Iliescu
27: 189451 Ion_Iliescu
28: 189451 Ion_Iliescu
29: 25445 Romania
30: 13275 Hungary
31: 189451 Ion_Iliescu
1: 189451 Ion_Iliescu
[pem_pos=0, final_scores=0.59330744,

```
final_local_scores=0.58023036, pem=0.0, similarity=21.404943]
2: 13275 Hungary
    [pem_pos=0, final_scores=0.61043745,
final_local_scores=0.6353366, pem=-0.32577527,
similarity=24.821041]
3: 36877 Bucharest
    [pem_pos=0, final_scores=0.4579665,
final_local_scores=0.46047717, pem=-0.01781679,
similarity=17.928162]
4: 25445 Romania
    [pem_pos=0, final_scores=0.7709686,
final_local_scores=0.7996473, pem=-0.21241479,
similarity=29.107927]
5: 189451 Ion_Iliescu
    [pem_pos=0, final_scores=0.7836091,
final_local_scores=0.788106, pem=0.0, similarity=27.608368]
6: 13275 Hungary
    [pem_pos=0, final_scores=0.52266407,
final_local_scores=0.53945744, pem=-0.32577527,
similarity=21.959816]
7: 189451 Ion_Iliescu
    [pem_pos=0, final_scores=0.60058004,
final_local_scores=0.58817464, pem=0.0, similarity=21.642017]
8: 407177 Social_Democratic_Party_(Romania)
    [pem_pos=0, final_scores=0.4621243,
final_local_scores=0.46002442, pem=0.0, similarity=17.817759]
9: 1969239 Romanian_National_Unity_Party
    [pem_pos=0, final_scores=0.54006976,
final_local_scores=0.5722278, pem=0.0, similarity=21.166132]
10: 1969239 Romanian_National_Unity_Party
    [pem_pos=0, final_scores=0.6986024,
final_local_scores=0.74540055, pem=0.0, similarity=26.333952]
11: 407177 Social_Democratic_Party_(Romania)
    [pem_pos=0, final_scores=0.47049862,
final_local_scores=0.46917212, pem=0.0, similarity=18.090744]
12: 395061 Adrian_Năstase
    [pem_pos=0, final_scores=0.83685094,
final_local_scores=0.87347, pem=0.0, similarity=30.155802]
13: 395061 Adrian_Năstase
    [pem_pos=0, final_scores=0.42153734,
final_local_scores=0.4198029, pem=0.0, similarity=16.61747]
14: 1969239 Romanian_National_Unity_Party
    [pem_pos=0, final_scores=0.617894,
final_local_scores=0.6572389, pem=0.0, similarity=23.703033]
15: 25445 Romania
    [pem_pos=0, final_scores=0.6179338,
final_local_scores=0.6324799, pem=-0.21241479,
similarity=24.11932]
16: 13275 Hungary
    [pem_pos=1, final_scores=0.6362687,
final_local_scores=0.66355324, pem=-1.0319586,
similarity=29.503416]
```

```
17: 21133 NATO
    [pem_pos=0, final_scores=0.3596266,
final_local_scores=0.4122005, pem=0.0, similarity=16.390598]
18: 9317 European_Union
    [pem_pos=0, final_scores=0.442325,
final_local_scores=0.48339438, pem=-0.005814796,
similarity=18.546787]
19: 1969239 Romanian_National_Unity_Party
    [pem_pos=0, final_scores=0.55480206,
final_local_scores=0.58832055, pem=0.0, similarity=21.646372]
20: 1969239 Romanian_National_Unity_Party
    [pem_pos=0, final_scores=0.5360504,
final_local_scores=0.5678372, pem=0.0, similarity=21.035107]
21: 1969239 Romanian_National_Unity_Party
    [pem_pos=0, final_scores=0.54819286,
final_local_scores=0.581101, pem=0.0, similarity=21.430925]
22: 1184425 Gheorghe_Funar
    [pem_pos=0, final_scores=0.5726334,
final_local_scores=0.621523, pem=0.0, similarity=22.6372]
23: 189451 Ion_Iliescu
    [pem_pos=0, final_scores=0.6307935,
final_local_scores=0.6211782, pem=0.0, similarity=22.62691]
24: 13275 Hungary
    [pem_pos=1, final_scores=0.6947189,
final_local_scores=0.7274012, pem=-1.0319586,
similarity=31.408766]
25: 1184425 Gheorghe_Funar
    [pem_pos=0, final_scores=0.48749715,
final_local_scores=0.5285246, pem=0.0, similarity=19.86194]
26: 189451 Ion_Iliescu
    [pem_pos=0, final_scores=0.5906754,
final_local_scores=0.57735527, pem=0.0, similarity=21.319145]
27: 189451 Ion_Iliescu
    [pem_pos=0, final_scores=0.55424577,
final_local_scores=0.5375615, pem=0.0, similarity=20.13162]
28: 189451 Ion_Iliescu
    [pem_pos=0, final_scores=0.6739554,
final_local_scores=0.668326, pem=0.0, similarity=24.033894]
29: 25445 Romania
    [pem_pos=1, final_scores=0.59663993,
final_local_scores=0.6092196, pem=-1.0382323,
similarity=27.916111]
30: 13275 Hungary
    [pem_pos=1, final_scores=0.6484656,
final_local_scores=0.6768765, pem=-1.0319586,
similarity=29.90101]
31: 189451 Ion_Iliescu
    [pem_pos=0, final_scores=0.6140292,
final_local_scores=0.60286576, pem=0.0, similarity=22.08043]
```

Sample: b'559_Estonian'

[1 1 Estonian 1] MPS see little hope of electing president .

Belinda Goldsmith [2 2 TALLINN 2] 1996-08-27 [3 3 Estonia 3] 's parliament failed for a second time to elect a president on Tuesday , dealing a blow to incumbent [4 4 Lennart Meri 4] and pushing the country towards stalemate in its choice of a new head of state . Neither [5 5 Meri 5] , who oversaw [6 6 Estonia 6] 's first steps into statehood after the collapse of the [7 7 Soviet Union 7] , nor his arch-rival , former communist [8 8 Arnold Ruutel 8] , have secured the 68 votes necessary from the 101-member parliament . [9 9 Meri 9] garnered 49 votes and [10 10 Meri 10] 34 in Tuesday 's ballot for the five-year presidency of [11 11 Estonia 11] . A third and final vote was due to be held when parliament reconvened on Tuesday but legislators were not expecting a clear result . If there is no result the decision will be ceded to an electoral college . " The votes are a strong message to [12 12 Meri 12] that he is not favoured by some politicians any more , " [13 13 Reform Party 13] head Heiki Kranich told [14 1 4 Reuters 14] . Under a constitution agreed in 1992 , a year after independence , the president has no executive powers . His only political role is to smooth the functioning of government in periods of crisis . But [15 15 Meri 15] , 67 , has been accused in parliament of taking too much power and not always consulting parliamentarians before making decisions . His relations with a leftist-led government have sometimes been tense . His support in the first round of voting on Monday was much lower than expected , scoring only 45 votes , which political analysts put down as a vote of no confidence in his performance . This support only inched up to 49 in the second vote . Support for [16 16 Ruutel 16] , 68 , remained constant at 34 votes . If the third vote fails to give either [17 17 Meri 17] or [18 18 Ruutel 18] 68 votes , the parliamentary speaker will convene an electoral college of 101 MPs and 273 local government representatives to hold a new poll that could include new nominations . This would be the first time that the former [19 19 Soviet 19] republic has had to call together an electoral college . In its first presidential election in 1992 [20 20 Meri 20] won the necessary votes in in a parliamentary election against [21 21 Ruutel 21] . Parliamentary organisers said the exact timetable remained unclear but it would probably take about a month to organise an electoral college which could also hold several rounds of voting before a clear winner emerges .

- 1: 28222445 Estonia
- 2: 31577 Tallinn
- 3: 28222445 Estonia
- 4: 379531 Lennart_Meri
- 5: 379531 Lennart_Meri
- 6: 28222445 Estonia
- 7: 26779 Soviet_Union
- 8: 341466 Arnold_Rüütel
- 9: 379531 Lennart_Meri
- 10: 379531 Lennart_Meri

```
11: 28222445 Estonia
12: 379531 Lennart_Meri
13: 424689 Estonian_Reform_Party
14: 18998750 Reuters
15: 379531 Lennart_Meri
16: 341466 Arnold_Rüütel
17: 379531 Lennart_Meri
18: 341466 Arnold_Rüütel
19: 26779 Soviet_Union
20: 379531 Lennart_Meri
21: 341466 Arnold_Rüütel
1: 28222445 Estonia
    [pem_pos=1, final_scores=0.5003882,
final_local_scores=0.5231766, pem=-0.96796864,
similarity=24.96631]
2: 31577 Tallinn
    [pem_pos=0, final_scores=0.4908021,
final_local_scores=0.5114361, pem=0.0, similarity=19.351986]
3: 28222445 Estonia
    [pem_pos=0, final_scores=0.60041714,
final_local_scores=0.63244295, pem=-0.4106629,
similarity=25.196321]
4: 379531 Lennart_Meri
    [pem_pos=0, final_scores=1.0557534,
final_local_scores=1.0429461, pem=0.0, similarity=35.213306]
5: 379531 Lennart_Meri
    [pem_pos=0, final_scores=0.98792917,
final_local_scores=0.9704643, pem=0.0, similarity=33.050304]
6: 28222445 Estonia
    [pem_pos=0, final_scores=0.5038633,
final_local_scores=0.5269726, pem=-0.4106629,
similarity=22.048876]
7: 26779 Soviet_Union
    [pem_pos=0, final_scores=0.4237187,
final_local_scores=0.43344802, pem=-0.075323656,
similarity=17.434288]
8: 341466 Arnold_Rüütel
    [pem_pos=0, final_scores=0.74508256,
final_local_scores=0.7811021, pem=0.0, similarity=27.399359]
9: 379531 Lennart_Meri
    [pem_pos=0, final_scores=0.91378075,
final_local_scores=0.8894684, pem=0.0, similarity=30.633224]
10: 379531 Lennart_Meri
    [pem_pos=0, final_scores=0.89476484,
final_local_scores=0.8686964, pem=0.0, similarity=30.013346]
11: 28222445 Estonia
    [pem_pos=0, final_scores=0.58445835,
final_local_scores=0.61501044, pem=-0.4106629,
similarity=24.6761]
12: 379531 Lennart_Meri
    [pem_pos=0, final_scores=0.92767876,
final_local_scores=0.9046499, pem=0.0, similarity=31.08627]
```

```
13: 424689 Estonian_Reform_Party
    [pem_pos=3, final_scores=0.33099043,
final_local_scores=0.38856, pem=-2.6619248,
similarity=30.16108]
14: 18998750 Reuters
    [pem_pos=0, final_scores=0.34724945,
final_local_scores=0.4373634, pem=0.0, similarity=17.141508]
15: 379531 Lennart_Meri
    [pem_pos=0, final_scores=0.92256397,
final_local_scores=0.8990628, pem=0.0, similarity=30.919542]
16: 341466 Arnold_Rüütel
    [pem_pos=0, final_scores=0.5553426,
final_local_scores=0.57384, pem=0.0, similarity=21.214243]
17: 379531 Lennart_Meri
    [pem_pos=0, final_scores=0.9306088,
final_local_scores=0.90785056, pem=0.0, similarity=31.181786]
18: 341466 Arnold_Rüütel
    [pem_pos=0, final_scores=0.597915,
final_local_scores=0.6203439, pem=0.0, similarity=22.602013]
19: 26779 Soviet_Union
    [pem_pos=0, final_scores=0.39040393,
final_local_scores=0.3970567, pem=-0.097171314,
similarity=16.46711]
20: 379531 Lennart_Meri
    [pem_pos=0, final_scores=0.9403505,
final_local_scores=0.9184918, pem=0.0, similarity=31.499338]
21: 341466 Arnold_Rüütel
    [pem_pos=0, final_scores=0.5346152,
final_local_scores=0.5511986, pem=0.0, similarity=20.538578]
```

Sample: b'560_Estonia'

[1 1 Estonia 1] assembly fails to elect state president . [2 2 TALLINN 2] 1996-08-27 The [3 3 Estonian 3] parliament failed for a third and final time to elect a new state president on Tuesday , refusing a second mandate for incumbent [4 4 Lennart Meri 4] . Neither [5 5 Meri 5] nor his rival [6 6 Arnold Ruutel 6] could garner the 68 votes needed from the 101 members of parliament to become president . In the third vote [7 7 Meri 7] won 52 and [8 8 Ruutel 8] won 32 votes . The final decision will now be made by a larger assembly . [9 9 Meri 9] won 49 in a second vote earlier on Tuesday and 45 in the first on Monday . [10 10 Ruutel 10] won 34 votes in the first two secret ballots . Enn Markwart , chairman of the National Election Commission said 96 members of parliament cast votes , with one ballot paper invalid and 11 abstentions . The election will now go before an electoral college involving MPs and local government representatives that will be convened by the parliamentary Speaker in the next day or so . It could take up to a month before a new vote but the timetable is not yet clear . This is the first time the former [11 11 Soviet 11] republic has had to convene such a group .

```
1: 28222445 Estonia
2: 31577 Tallinn
3: 28222445 Estonia
4: 379531 Lennart_Meri
5: 379531 Lennart_Meri
6: 341466 Arnold_Rüütel
7: 379531 Lennart_Meri
8: 341466 Arnold_Rüütel
9: 379531 Lennart_Meri
10: 341466 Arnold_Rüütel
11: 26779 Soviet_Union
1: 28222445 Estonia
    [pem_pos=0, final_scores=0.60544676,
final_local_scores=0.6422798, pem=-0.4106629,
similarity=25.489872]
2: 31577 Tallinn
    [pem_pos=0, final_scores=0.49801862,
final_local_scores=0.520911, pem=0.0, similarity=19.634735]
3: 28222445 Estonia
    [pem_pos=1, final_scores=0.47050595,
final_local_scores=0.49487746, pem=-0.96796864,
similarity=24.121807]
4: 379531 Lennart_Meri
    [pem_pos=0, final_scores=1.0661896,
final_local_scores=1.0615084, pem=0.0, similarity=35.767242]
5: 379531 Lennart_Meri
    [pem_pos=0, final_scores=0.94842094,
final_local_scores=0.9377925, pem=0.0, similarity=32.07531]
6: 341466 Arnold_Rüütel
    [pem_pos=0, final_scores=0.7604676,
final_local_scores=0.79139143, pem=0.0, similarity=27.706411]
7: 379531 Lennart_Meri
    [pem_pos=0, final_scores=1.0022403,
final_local_scores=0.996582, pem=0.0, similarity=33.82971]
8: 341466 Arnold_Rüütel
    [pem_pos=0, final_scores=0.6603387,
final_local_scores=0.68201584, pem=0.0, similarity=24.442427]
9: 379531 Lennart_Meri
    [pem_pos=0, final_scores=0.84071654,
final_local_scores=0.82014173, pem=0.0, similarity=28.564379]
10: 341466 Arnold_Rüütel
    [pem_pos=0, final_scores=0.46391636,
final_local_scores=0.4674543, pem=0.0, similarity=18.039482]
11: 26779 Soviet_Union
    [pem_pos=0, final_scores=0.40757638,
final_local_scores=0.42574483, pem=-0.097171314,
similarity=17.323221]
```

Sample: b'561_Albania'
[1 1 Albania 1] charges [2 2 Briton 2] with child sex
abuse . [3 3 TIRANA 3] 1996-08-27 [4 4 Albanian 4]
authorities have arrested and charged a [5 5 British 5] man

for sexually abusing two young boys , a [6 6 Tirana 6] prosecutor said on Tuesday . " We have arrested him and charged him with these shameful acts of sex abuse of little children , " prosecutor Adnan Xhelili told [7 7 Reuters 7] . Xhelili said Paul Thompson , 34 , from [8 8 Wiltshire 8] , was arrested on Sunday in a hotel in the [9 9 Adriatic 9] resort of [10 10 Durres 10] , 45 km (30 miles) west of [11 1 1 Tirana 11] . Thompson has denied the charges . He could face up to five years in jail if convicted . Xhelili said Thompson , who is divorced , said he befriended the boys , both aged under 10 , because they reminded him of his own children who live with his former wife in [12 12 London 12] . The prosecutor 's office said no date had yet been set for a trial to begin as investigations had first to be completed . The [13 13 British 13] embassy in [14 14 Tirana 14] said it had sent an embassy official to talk to Thompson who is being held in jail . The age of consent for heterosexual and homosexual sex in [15 15 Albania 15] is 14 . A large number of destitute children can be seen begging in the streets of impoverished [16 16 Albania 16] , especially in towns and resorts visited by foreigners .

1: 738 Albania
2: 31717 United_Kingdom
3: 31075 Tirana
4: 738 Albania
5: 31717 United_Kingdom
6: 31075 Tirana
7: 18998750 Reuters
8: 51231 Wiltshire
9: 23275478 Adriatic_Sea
10: 263530 Durrës
11: 31075 Tirana
12: 17867 London
13: 31717 United_Kingdom
14: 31075 Tirana
15: 738 Albania
16: 738 Albania
1: 738 Albania
[pem_pos=0, final_scores=0.6085107,
final_local_scores=0.5738772, pem=-0.41162592,
similarity=23.453838]
2: 31717 United_Kingdom
[pem_pos=1, final_scores=0.35424513,
final_local_scores=0.3201119, pem=-0.9694005,
similarity=18.914242]
3: 31075 Tirana
[pem_pos=0, final_scores=0.42778617,
final_local_scores=0.39161968, pem=0.0, similarity=15.776425]
4: 738 Albania
[pem_pos=2, final_scores=0.56907505,
final_local_scores=0.5307998, pem=-1.5560378,
similarity=28.391813]

```
5: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.52232116,
final_local_scores=0.50370944, pem=-0.48936236,
similarity=21.782637]
6: 31075 Tirana
    [pem_pos=0, final_scores=0.38623518,
final_local_scores=0.34623158, pem=0.0, similarity=14.421954]
7: 18998750 Reuters
    [pem_pos=0, final_scores=0.35902184,
final_local_scores=0.40689027, pem=0.0, similarity=16.23213]
8: 51231 Wiltshire
    [pem_pos=0, final_scores=0.47472256,
final_local_scores=0.53739667, pem=-0.030204631,
similarity=20.290958]
9: 23275478 Adriatic_Sea
    [pem_pos=0, final_scores=0.36579287,
final_local_scores=0.39642864, pem=0.0, similarity=15.919934]
10: 263530 Durrës
    [pem_pos=0, final_scores=0.4426914,
final_local_scores=0.4604463, pem=0.0, similarity=17.830349]
11: 31075 Tirana
    [pem_pos=0, final_scores=0.41505307,
final_local_scores=0.3777107, pem=0.0, similarity=15.361353]
12: 17867 London
    [pem_pos=0, final_scores=0.5348942,
final_local_scores=0.53728724, pem=-0.09086737,
similarity=20.617586]
13: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.608689,
final_local_scores=0.6072856, pem=0.0, similarity=22.212328]
14: 31075 Tirana
    [pem_pos=0, final_scores=0.40785342,
final_local_scores=0.36984622, pem=0.0, similarity=15.126661]
15: 738 Albania
    [pem_pos=0, final_scores=0.562787,
final_local_scores=0.523931, pem=-0.41162592,
similarity=21.963345]
16: 738 Albania
    [pem_pos=0, final_scores=0.52632016,
final_local_scores=0.48409653, pem=-0.41162592,
similarity=20.774605]
```

Sample: b'562_Estonia'

```
[1 1 Estonia 1] assembly again fails to elect president . [2
2 TALLINN 2] 1996-08-27 [3 3 Estonia 3] 's parliament again
failed to elect a new state president on Tuesday when neither
of two candidates secured a majority in second-round voting .
Incumbent president [4 4 Lennart Meri 4] won 49 votes
compared to 34 won by his rival , deputy Parliamentary
Speaker [5 5 Arnold Ruutel 5] . But [6 6 Meri 6] 's support
was not enough for the 68 needed for election and a third
secret ballot will take place later in the day ( 1300 [7 7 GMT
```

7]) , parliamentary officials said . To win a clear mandate for the five-year presidential term , a candidate must secure 68 votes from the 101-member parliament . Enn Markwart , Chairman of the National Election Commission , said 96 members of parliament voted in the second round , with 12 abstentions and one ballot paper invalid . On Monday , in the first round of voting , [8 8 Meri 8] secured 45 votes and [9 9 Ruutel 9] 34 . [10 10 Meri 10] 's popularity has suffered in recent years , with politicians criticising him for taking too much power and acting without consulting parliament . If a third round of voting fails to give either candidate 68 votes , the parliamentary speaker has to convene an electoral college of all 101 MPs and 273 local government representatives for a new vote that could take up to a month .

1: 28222445 Estonia
2: 31577 Tallinn
3: 28222445 Estonia
4: 379531 Lennart_Meri
5: 341466 Arnold_Rüütel
6: 379531 Lennart_Meri
7: 12701 Greenwich_Mean_Time
8: 379531 Lennart_Meri
9: 341466 Arnold_Rüütel
10: 379531 Lennart_Meri
1: 28222445 Estonia
[pem_pos=0, final_scores=0.5755839,
final_local_scores=0.6081328, pem=-0.4106629,
similarity=24.470858]
2: 31577 Tallinn
[pem_pos=0, final_scores=0.48217648,
final_local_scores=0.50308377, pem=0.0, similarity=19.102736]
3: 28222445 Estonia
[pem_pos=0, final_scores=0.54544467,
final_local_scores=0.5752103, pem=-0.4106629,
similarity=23.488384]
4: 379531 Lennart_Meri
[pem_pos=0, final_scores=1.0363094,
final_local_scores=1.0258389, pem=0.0, similarity=34.702793]
5: 341466 Arnold_Rüütel
[pem_pos=0, final_scores=0.8723791,
final_local_scores=0.93245596, pem=0.0, similarity=31.916058]
6: 379531 Lennart_Meri
[pem_pos=0, final_scores=0.9014289,
final_local_scores=0.8833963, pem=0.0, similarity=30.45202]
7: 12701 Greenwich_Mean_Time
[pem_pos=0, final_scores=0.601301,
final_local_scores=0.7570304, pem=0.0, similarity=26.681011]
8: 379531 Lennart_Meri
[pem_pos=0, final_scores=1.0082897,
final_local_scores=1.0001256, pem=0.0, similarity=33.935455]
9: 341466 Arnold_Rüütel
[pem_pos=0, final_scores=0.61196524,

```
final_local_scores=0.64799327, pem=0.0, similarity=23.427124]
10: 379531 Lennart_Meri
    [pem_pos=0, final_scores=0.8734217,
final_local_scores=0.8528028, pem=0.0, similarity=29.53905]
```

Sample: b'563_Slovak'

[1 1 Slovak 1] women visited [2 2 Dutroux 2] , police say . Peter Laca [3 3 BRATISLAVA 3] 1996-08-27 [4 4 Marc Dutroux 4] , the chief accused in [5 5 Belgium 5] 's sensational child murder and sex abuse case , visited [6 6 Slovakia 6] a number of times and about 10 young [7 7 Slovak 7] women went to [8 8 Belgium 8] at his invitation , police said on Tuesday . But they have difficulty remembering what happened there , perhaps because of drugs , and are unsure whether they were filmed for pornography , Rudolf Gajdos , head of the [9 9 Slo vak 9] office of Interpol , told [10 10 Reuters 10] . Although Gajdos spoke of " girls " his deputy , Eva Boudova , said the case involved about 10 young women in their early 20s . " The police interrogated several [11 11 Slovak 11] girls who said that they had been invited by [12 12 Mark Dutr oux 12] to visit [13 13 Belgium 13] , " Gajdos said . " The girls said they went to [14 14 Belgium 14] voluntarily and the police suspect that they were used to act in pornographic films . " " The police suspect (the girls) were under the influence of drugs as some girls admitted they took unspecified pills . " " We have suspicions of a rape , but the police still have to find the victim , " Gajdos added . [1 5 15 Dutroux 15] 's last visit to [16 16 Slovakia 16] was reported to have been as recent as July . [17 17 Slovak 17] police are also cooperating with [18 18 Belgium 18] in the search for An Marchal and Eefje Lambrecks , who went missing last August . [19 19 Dutroux 19] , 39 , who was charged last week with the abduction and illegal imprisonment of two other girls aged 14 and 12 , is one of several suspects in the Marchal and Lambrecks case . Last Saturday he led police to the bodies of two other girls , aged eight , who died of starvation this year after their abduction in June , 1995 . The [20 20 Czech 20] office of Interpol said on Friday it would neither confirm nor deny that [21 21 Dutroux 21] had been in the [22 22 Czech Republic 22] , [23 23 Slovakia 23] 's western neighbour and former federation partner . [24 24 Be lgian 24] police said an officer had visited [25 25 Bratislava 25] to talk with [26 26 Slovak 26] detectives about An and Eefje and other disappearances , and he planned to go also to [27 27 Prague 27] .

1: 26830 Slovakia
2: 154694 Marc_Dutroux
3: 18933194 Bratislava
4: 154694 Marc_Dutroux
5: 3343 Belgium
6: 26830 Slovakia
7: 26830 Slovakia

```
8: 3343 Belgium
9: 26830 Slovakia
10: 18998750 Reuters
11: 26830 Slovakia
12: 154694 Marc_Dutroux
13: 3343 Belgium
14: 3343 Belgium
15: 154694 Marc_Dutroux
16: 26830 Slovakia
17: 26830 Slovakia
18: 3343 Belgium
19: 154694 Marc_Dutroux
20: 5321 Czech_Republic
21: 154694 Marc_Dutroux
22: 5321 Czech_Republic
23: 26830 Slovakia
24: 3343 Belgium
25: 18933194 Bratislava
26: 26830 Slovakia
27: 23844 Prague
1: 26830 Slovakia
    [pem_pos=1, final_scores=0.65988123,
final_local_scores=0.62850255, pem=-1.0527612,
similarity=28.570562]
2: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.60110366,
final_local_scores=0.6820591, pem=0.0, similarity=24.443718]
3: 18933194 Bratislava
    [pem_pos=0, final_scores=0.44192594,
final_local_scores=0.41343188, pem=0.0, similarity=16.427345]
4: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.60296524,
final_local_scores=0.6840926, pem=0.0, similarity=24.5044]
5: 3343 Belgium
    [pem_pos=0, final_scores=0.6018151,
final_local_scores=0.56724805, pem=-0.24180195,
similarity=22.332483]
6: 26830 Slovakia
    [pem_pos=0, final_scores=0.58487505,
final_local_scores=0.54656965, pem=-0.36602598,
similarity=22.390947]
7: 26830 Slovakia
    [pem_pos=1, final_scores=0.6243512,
final_local_scores=0.58969134, pem=-1.0527612,
similarity=27.41236]
8: 3343 Belgium
    [pem_pos=0, final_scores=0.57254714,
final_local_scores=0.5352772, pem=-0.24180195,
similarity=21.378408]
9: 26830 Slovakia
    [pem_pos=1, final_scores=0.6866856,
final_local_scores=0.65778226, pem=-1.0527612,
```

```
similarity=29.444328]
10: 18998750 Reuters
    [pem_pos=0, final_scores=0.36824828,
final_local_scores=0.45597094, pem=0.0, similarity=17.696795]
11: 26830 Slovakia
    [pem_pos=1, final_scores=0.70252603,
final_local_scores=0.6750855, pem=-1.0527612,
similarity=29.960688]
12: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.69472104,
final_local_scores=0.7843219, pem=0.0, similarity=27.495443]
13: 3343 Belgium
    [pem_pos=0, final_scores=0.546159,
final_local_scores=0.5064522, pem=-0.24180195,
similarity=20.518213]
14: 3343 Belgium
    [pem_pos=0, final_scores=0.6104403,
final_local_scores=0.57666975, pem=-0.24180195,
similarity=22.613646]
15: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.541485,
final_local_scores=0.6169348, pem=0.0, similarity=22.500277]
16: 26830 Slovakia
    [pem_pos=0, final_scores=0.6044488,
final_local_scores=0.567951, pem=-0.36602598,
similarity=23.02901]
17: 26830 Slovakia
    [pem_pos=1, final_scores=0.69615144,
final_local_scores=0.66812223, pem=-1.0527612,
similarity=29.752893]
18: 3343 Belgium
    [pem_pos=0, final_scores=0.5572122,
final_local_scores=0.51852614, pem=-0.24180195,
similarity=20.878525]
19: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.53558373,
final_local_scores=0.61048853, pem=0.0, similarity=22.307909]
20: 5321 Czech_Republic
    [pem_pos=1, final_scores=0.46787864,
final_local_scores=0.41884297, pem=-1.208016,
similarity=23.158201]
21: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.643912,
final_local_scores=0.72882074, pem=0.0, similarity=25.839178]
22: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.47112983,
final_local_scores=0.42239445, pem=-0.27794555,
similarity=18.206318]
23: 26830 Slovakia
    [pem_pos=0, final_scores=0.650355,
final_local_scores=0.61809653, pem=-0.36602598,
similarity=24.525454]
```

```
24: 3343 Belgium
    [pem_pos=0, final_scores=0.66039413,
final_local_scores=0.63123673, pem=-0.12403572,
similarity=23.6016]
25: 18933194 Bratislava
    [pem_pos=0, final_scores=0.5738593,
final_local_scores=0.55754906, pem=0.0, similarity=20.728088]
26: 26830 Slovakia
    [pem_pos=1, final_scores=0.6524863,
final_local_scores=0.6204247, pem=-1.0527612,
similarity=28.329504]
27: 23844 Prague
    [pem_pos=0, final_scores=0.45439798,
final_local_scores=0.44581687, pem=-0.05392049,
similarity=17.687006]
```

Sample: b'564_WEATHER'

WEATHER - Conditions at CIS airports - August 27 . [1 1 MOSCOW 1] 1996-08-27 No closures of airports due to bad weather are expected in the Commonwealth of Independent States on August 28 and August 29 , the Russian Weather Service said on Tuesday . -- Moscow Newsroom +7095 941 8520

```
1: 19004 Moscow
1: 19004 Moscow
    [pem_pos=0, final_scores=0.39234346,
final_local_scores=0.7011628, pem=-0.106582455,
similarity=25.593422]
```

Sample: b'565_Russian'

[1 1 Russian 1] serial killer strikes again . [2 2 MOSCOW 2] 1996-08-27 [3 3 Russian 3] police in the [4 4 Urals 4] city of [5 5 Perm 5] are on the trail of a serial killer who has claimed his seventh victim in just a few months , [6 6 Itar-Tass 6] news agency said on Tuesday . In the latest attack , the killer raped and stabbed a young woman in a lift , leaving her body on a landing . Tass did not say exactly when it took place . Police earlier released a suspect after women who had survived an attack failed to identify him .

```
1: 25391 Russia
2: 19004 Moscow
3: 25391 Russia
4: 32152 Ural_Mountains
5: 389777 Perm
6: 434852 Information_Telegraph_Agency_of_Russia
1: 25391 Russia
    [pem_pos=1, final_scores=0.92368716,
final_local_scores=0.91603786, pem=-1.0515022,
similarity=37.144344]
2: 19004 Moscow
    [pem_pos=0, final_scores=0.6749212,
final_local_scores=0.67206377, pem=-0.106582455,
similarity=24.725048]
```

```
3: 25391 Russia
    [pem_pos=1, final_scores=0.9687225,
final_local_scores=0.9652322, pem=-1.0515022,
similarity=38.6124]
4: 32152 Ural_Mountains
    [pem_pos=0, final_scores=0.4141081,
final_local_scores=0.42747885, pem=0.0, similarity=16.846535]
5: 389777 Perm
    [pem_pos=0, final_scores=0.4606437,
final_local_scores=0.44474506, pem=-0.0032467444,
similarity=17.379448]
6: 434852 Information_Telegraph_Agency_of_Russia
    [pem_pos=0, final_scores=0.7306385,
final_local_scores=0.84434384, pem=0.0, similarity=29.286617]
```

Sample: b'566_Russian'

[1 1 Russian 1] army , [2 2 Chechens 2] open new round of talks . [3 3 MOSCOW 3] 1996-08-27 [4 4 Russia 4] 's military commander in [5 5 Chechnya 5] began new talks with separatist chief-of-staff [6 6 Aslan Maskhadov 6] on Tuesday , [7 7 Itar-Tass 7] news agency said . Tass said the talks were taking place in the settlement of [8 8 Novye Atagi 8] , some 20 km (12 miles) south of the [9 9 Chechen 9] capital [10 10 Grozny 10] . The talks had been postponed while the [11 11 Russia 11] waited for the rebels to return arms and ammunition seized from [12 12 Russian 12] soldiers at the weekend . The [13 13 Chechens 13] said on Monday they had returned all the weapons , which they said were seized by a renegade group . The talks between [14 14 Maskhadov 14] and [15 15 Russia 15] 's Vyacheslav Tikhomirov are aimed at putting the finishing touches to a ceasefire sealed last week in talks with [16 16 Russian 16] security chief [17 17 Alexander Lebed 17] . [18 18 Lebed 18] , who met [19 19 Russian 19] Prime Minister [20 20 Viktor Chernomyrdin 20] on Monday to discuss the progress he made on a political settlement for the breakaway region , has been seeking a meeting with President [21 21 Boris Yeltsin 21] , who started a holiday near [22 22 Moscow 22] on Monday .

```
1: 25391 Russia
2: 6095 Chechnya
3: 19004 Moscow
4: 25391 Russia
5: 6095 Chechnya
6: 340963 Aslan_Maskhadov
7: 434852 Information_Telegraph_Agency_of_Russia
8: 4813770 Novye_Atagi
9: 6095 Chechnya
10: 140790 Grozny
11: 25391 Russia
12: 25391 Russia
13: 6095 Chechnya
14: 340963 Aslan_Maskhadov
15: 25391 Russia
```

```
16: 25391 Russia
17: 431579 Alexander_Lebed
18: 431579 Alexander_Lebed
19: 25391 Russia
20: 482420 Viktor_Chernomyrdin
21: 23278520 Boris_Yeltsin
22: 19004 Moscow
1: 25391 Russia
    [pem_pos=1, final_scores=0.92430097,
final_local_scores=0.93495816, pem=-1.0515022,
similarity=37.70896]
2: 6095 Chechnya
    [pem_pos=2, final_scores=0.47581607,
final_local_scores=0.43745726, pem=-0.96860826,
similarity=22.411753]
3: 19004 Moscow
    [pem_pos=0, final_scores=0.64344233,
final_local_scores=0.6392144, pem=-0.106582455,
similarity=23.744759]
4: 25391 Russia
    [pem_pos=0, final_scores=0.8549133,
final_local_scores=0.85916275, pem=-0.20299491,
similarity=30.832762]
5: 6095 Chechnya
    [pem_pos=0, final_scores=0.5513056,
final_local_scores=0.5199181, pem=0.0, similarity=19.605106]
6: 340963 Aslan_Maskhadov
    [pem_pos=0, final_scores=0.607611,
final_local_scores=0.62878126, pem=0.0, similarity=22.8538]
7: 434852 Information_Telegraph_Agency_of_Russia
    [pem_pos=0, final_scores=0.728192,
final_local_scores=0.80028844, pem=0.0, similarity=27.971916]
8: 4813770 Novye_Atagi
    [pem_pos=0, final_scores=0.5294651,
final_local_scores=0.6288606, pem=0.0, similarity=22.856167]
9: 6095 Chechnya
    [pem_pos=2, final_scores=0.5002988,
final_local_scores=0.4642009, pem=-0.8617233,
similarity=22.62858]
10: 140790 Grozny
    [pem_pos=0, final_scores=0.5067386,
final_local_scores=0.49357283, pem=0.0, similarity=18.81891]
11: 25391 Russia
    [pem_pos=1, final_scores=0.66927403,
final_local_scores=0.65638, pem=-1.8802736,
similarity=33.90262]
12: 25391 Russia
    [pem_pos=1, final_scores=0.91124684,
final_local_scores=0.9206986, pem=-1.0515022,
similarity=37.28343]
13: 6095 Chechnya
    [pem_pos=2, final_scores=0.48453003,
```

```
final_local_scores=0.44697595, pem=-0.96860826,
similarity=22.695808]
14: 340963 Aslan_Maskhadov
    [pem_pos=0, final_scores=0.4733277,
final_local_scores=0.4820972, pem=0.0, similarity=18.476456]
15: 25391 Russia
    [pem_pos=0, final_scores=0.9657677,
final_local_scores=0.98025435, pem=-0.20299491,
similarity=34.446377]
16: 25391 Russia
    [pem_pos=1, final_scores=0.92113835,
final_local_scores=0.9315035, pem=-1.0515022,
similarity=37.60587]
17: 431579 Alexander_Lebed
    [pem_pos=0, final_scores=0.7578972,
final_local_scores=0.7951708, pem=0.0, similarity=27.819195]
18: 431579 Alexander_Lebed
    [pem_pos=0, final_scores=0.6730722,
final_local_scores=0.7122997, pem=0.0, similarity=25.346157]
19: 25391 Russia
    [pem_pos=1, final_scores=0.9495843,
final_local_scores=0.96257645, pem=-1.0515022,
similarity=38.53315]
20: 482420 Viktor_Chernomyrdin
    [pem_pos=0, final_scores=0.7663979,
final_local_scores=0.828202, pem=0.0, similarity=28.804913]
21: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.6439109,
final_local_scores=0.6528518, pem=0.0, similarity=23.572113]
22: 19004 Moscow
    [pem_pos=0, final_scores=0.6667002,
final_local_scores=0.66462, pem=-0.106582455,
similarity=24.50291]
```

Sample: b'567_Argentine'

[1 1 Argentine 1] bishop reminds cabinet of Commandments . [2 2 BUENOS AIRES 2] , [3 3 Argentina 3] 1996-08-27 The Archbishop of [4 4 Buenos Aires 4] said on Tuesday the first thing he would do if elected president of [5 5 Argentina 5] would be to put up posters of the [6 6 Ten Commandments 6] in government offices . " They asked me what would be the first thing I would do if I were president , and I said the first thing I would do would be to resign straight away , " Archbishop [7 7 Antonio Quarracino 7] said at a sermon attended by several cabinet ministers . " But before going , I would have big signs put up in all government offices , those to do with justice , in all sectors , with the [8 8 Ten Commandments 8] , " he added . [9 9 Argentina 9] 's top [10 10 Roman Catholic 10] cleric said the [11 11 Biblical 11] commandment " Thou shalt not steal " would get special emphasis " because it has to be about the most common thing these days . " [12 12 Quarracino 12] and other Church leaders

are regular critics of the government 's free-market economic policy .

1: 18951905 Argentina
2: 255919 Buenos_Aires
3: 18951905 Argentina
4: 255919 Buenos_Aires
5: 18951905 Argentina
6: 2539671 Ten_Commandments
7: 5007152 Antonio_Quarracino
8: 2539671 Ten_Commandments
9: 18951905 Argentina
10: 606848 Catholic_Church
11: 3390 Bible
12: 5007152 Antonio_Quarracino
1: 18951905 Argentina
[pem_pos=0, final_scores=0.66024095,
final_local_scores=0.65519357, pem=-0.18885723,
similarity=24.669031]
2: 255919 Buenos_Aires
[pem_pos=0, final_scores=0.434735,
final_local_scores=0.41073692, pem=-0.12597719,
similarity=17.032005]
3: 18951905 Argentina
[pem_pos=0, final_scores=0.6103117,
final_local_scores=0.6006535, pem=-0.30003807,
similarity=23.646065]
4: 255919 Buenos_Aires
[pem_pos=0, final_scores=0.38198978,
final_local_scores=0.35312074, pem=-0.12597719,
similarity=15.312623]
5: 18951905 Argentina
[pem_pos=0, final_scores=0.51464987,
final_local_scores=0.49615747, pem=-0.30003807,
similarity=20.527695]
6: 2539671 Ten_Commandments
[pem_pos=0, final_scores=0.3607177,
final_local_scores=0.44190222, pem=0.0, similarity=17.276957]
7: 5007152 Antonio_Quarracino
[pem_pos=0, final_scores=0.5459197,
final_local_scores=0.61475086, pem=0.0, similarity=22.435104]
8: 2539671 Ten_Commandments
[pem_pos=0, final_scores=0.3982172,
final_local_scores=0.48286474, pem=0.0, similarity=18.49936]
9: 18951905 Argentina
[pem_pos=0, final_scores=0.5577355,
final_local_scores=0.543222, pem=-0.30003807,
similarity=21.932194]
10: 606848 Catholic_Church
[pem_pos=0, final_scores=0.39319247,
final_local_scores=0.4012789, pem=0.0, similarity=16.064676]
11: 3390 Bible
[pem_pos=0, final_scores=0.29194885,

```
final_local_scores=0.32653946, pem=-0.23351274,
similarity=15.104181]
12: 5007152 Antonio_Quarracino
    [pem_pos=0, final_scores=0.36438334,
final_local_scores=0.41644996, pem=0.0, similarity=16.51741]

Sample: b'568_Brazil'
[1 1 Brazil 1] 's [2 2 Eletropaulo 2] names new president . [
3 3 SAO PAULO 3] 1996-08-27 [4 4 Sao Paulo 4] state power
firm [5 5 Eletropaulo 5] said it has named Eduardo Bernini as
new president , replacing Emmanuel Sobral , who will head a
secretariat at the Transportation Ministry . Bernini is
expected to take office Thursday , a [6 6 Eletropaulo 6]
spokeswoman said . -- Romina Nicaretta , [7 7 Sao Paulo 7]
newsroom , 5511 232 4411 .
1: 3383 Brazil
2: 4315911 AES_Eletropaulo
3: 390875 São_Paulo
4: 390875 São_Paulo
5: 4315911 AES_Eletropaulo
6: 4315911 AES_Eletropaulo
7: 390875 São_Paulo
1: 3383 Brazil
    [pem_pos=0, final_scores=0.715035,
final_local_scores=0.764866, pem=-0.21313342,
similarity=28.073893]
2: 4315911 AES_Eletropaulo
    [pem_pos=0, final_scores=0.49642813,
final_local_scores=0.521862, pem=0.0, similarity=19.663115]
3: 390875 São_Paulo
    [pem_pos=0, final_scores=0.42257053,
final_local_scores=0.40340173, pem=0.0, similarity=16.128025]
4: 390875 São_Paulo
    [pem_pos=0, final_scores=0.23128013,
final_local_scores=0.194446, pem=0.0, similarity=9.892369]
5: 4315911 AES_Eletropaulo
    [pem_pos=0, final_scores=0.53592366,
final_local_scores=0.5650049, pem=0.0, similarity=20.950586]
6: 4315911 AES_Eletropaulo
    [pem_pos=0, final_scores=0.48776633,
final_local_scores=0.51240027, pem=0.0, similarity=19.380758]
7: 390875 São_Paulo
    [pem_pos=0, final_scores=0.35079145,
final_local_scores=0.32499397, pem=0.0, similarity=13.788181]

Sample: b'569_Dutch'
[1 1 Dutch 1] government wo n't pay ransom for kidnap
victims . [2 2 SAN JOSE 2] , [3 3 Costa Rica 3] 1996-08-27 [4
4 The Netherlands 4] government has ruled out paying ransom
money for a [5 5 Dutch 5] couple kidnapped from their farm ,
while [6 6 Costa Rican 6] authorities said on Tuesday they
had no leads in the case . " We have not had contact with the
```

kidnappers nor do we have any leads to take us to where they might be held , " Chief of Judicial Police Manuel Alvarado told [7 7 Reuters 7] . Hurte Sierd Zylstra and his wife , Jetsi Hendrika Coers , both 50 , were abducted from the teak plantation they managed late Saturday or early Sunday by at least two men demanding \$ 1.5 million ransom , authorities said . [8 8 Costa Rican 8] officials on Monday had given different names for the couple . Anton Schutte , an official with the embassy for [9 9 Belgium 9] , the [10 10 Netherlands 10] and [11 11 Luxembourg 11] , said the [12 12 Dutch 12] government had ruled out paying any money in ransom . " We 're looking at a criminal act that has no political aspect as far as what we can tell , " Schutte added . A note with the ransom demand was left in the couple 's car , which was used in the kidnapping , Schutte told a news conference on Monday . He said the note , believed to have been hand-written in [13 13 Spanish 13] and signed by the victims , was addressed to Ebe Huizinga , another [14 14 Dutch 14] citizen who owns the tree plantation . " Depending on you , we will either live or die , " it said . Alvarado said the car was abandoned about 40 miles (60 km) north of the couple 's house but said that did not indicate the kidnappers intended to take their victims into neighbouring [15 15 Nicaragua 15] . The farm is in the border region where a [16 16 German 1 6] tourist and a [17 17 Swiss 17] tour guide were kidnapped last New Year 's Eve by a heavily armed band led by a former [18 18 Nicaraguan 18] guerrilla . The two were held for 71 days until relatives paid a ransom . Two of the suspected abductors have since been arrested .

1: 21148 Netherlands
2: 219665 San_José,_Costa_Rica
3: 5551 Costa_Rica
4: 21148 Netherlands
5: 21148 Netherlands
6: 5551 Costa_Rica
7: 18998750 Reuters
8: 5551 Costa_Rica
9: 3343 Belgium
10: 21148 Netherlands
11: 17515 Luxembourg
12: 21148 Netherlands
13: 26825 Spanish_language
14: 21148 Netherlands
15: 21362 Nicaragua
16: 11867 Germany
17: 26748 Switzerland
18: 21362 Nicaragua
1: 21148 Netherlands
[pem_pos=0, final_scores=0.6548778,
final_local_scores=0.6206997, pem=-0.7943452,
similarity=26.932404]
2: 219665 San_José,_Costa_Rica

```
[pem_pos=1, final_scores=0.56405073,
final_local_scores=0.5832005, pem=-2.9430156,
similarity=37.498158]
3: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.54482895,
final_local_scores=0.5628265, pem=-0.30652362,
similarity=22.552502]
4: 21148 Netherlands
    [pem_pos=0, final_scores=0.66169953,
final_local_scores=0.6281515, pem=-0.10774545,
similarity=23.420942]
5: 21148 Netherlands
    [pem_pos=0, final_scores=0.5947343,
final_local_scores=0.5550021, pem=-0.7943452,
similarity=24.971855]
6: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.6183025,
final_local_scores=0.64308524, pem=0.0, similarity=23.280659]
7: 18998750 Reuters
    [pem_pos=0, final_scores=0.31294155,
final_local_scores=0.37840956, pem=0.0, similarity=15.382209]
8: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.57220083,
final_local_scores=0.59272623, pem=0.0, similarity=21.777845]
9: 3343 Belgium
    [pem_pos=0, final_scores=0.5411833,
final_local_scores=0.5025696, pem=-0.24180195,
similarity=20.40235]
10: 21148 Netherlands
    [pem_pos=0, final_scores=0.47158015,
final_local_scores=0.4204749, pem=-0.24940795,
similarity=17.993841]
11: 17515 Luxembourg
    [pem_pos=0, final_scores=0.49306607,
final_local_scores=0.4675184, pem=-0.49845144,
similarity=20.75205]
12: 21148 Netherlands
    [pem_pos=0, final_scores=0.6390732,
final_local_scores=0.6034356, pem=-0.7943452,
similarity=26.417208]
13: 26825 Spanish_language
    [pem_pos=0, final_scores=0.30424732,
final_local_scores=0.31986994, pem=-0.6062411,
similarity=16.932104]
14: 21148 Netherlands
    [pem_pos=0, final_scores=0.616051,
final_local_scores=0.57828736, pem=-0.7943452,
similarity=25.666735]
15: 21362 Nicaragua
    [pem_pos=0, final_scores=0.5031443,
final_local_scores=0.49834073, pem=-0.24729352,
similarity=20.306015]
```

```
16: 11867 Germany
    [pem_pos=1, final_scores=0.74395233,
final_local_scores=0.73873, pem=-1.0465823,
similarity=31.826368]
17: 26748 Switzerland
    [pem_pos=0, final_scores=0.6994677,
final_local_scores=0.69090796, pem=-0.1392413,
similarity=25.465]
18: 21362 Nicaragua
    [pem_pos=0, final_scores=0.45367146,
final_local_scores=0.44429916, pem=0.0, similarity=17.348486]
```

Sample: b'570_Venezuela'

[1 1 Venezuela 1] unions harden towards [2 2 CVG 2] privatization . [3 3 CARACAS 3] 1996-08-27 A swell of protest is growing within [4 4 Venezuela 4] 's trade unions at the proposed year-end privatization of the state-owned holding company [5 5 Corporacion Venezolana de Guayana 5] ([6 6 CVG 6]) , [7 7 CVG 7] union leaders said Tuesday . " We oppose the way the government is proceeding with the sale , " Ramon Machuca , [8 8 Sidor 8] trade union Secretary General and member of union-based opposition party [9 9 Radical Cause 9] , told reporters . " We do n't believe the government will make its timetable , " he added . [10 10 Sidor 10] is the [11 11 CVG 11] 's steel-producing arm , slated for a December sale worth an estimated \$ 1.5 billion . The [12 12 CVG 12] 's aluminum companies Venalum and Alucasa are also scheduled to be sold early 1997 . Arguing that [13 13 CVG 13] 's privatization would result in some 13,000 layoffs , compared to the government 's estimated 1,500 , [14 14 CVG 14] 's union leaders told reporters they would strike and stage protests if their concerns were not addressed . " We oppose any privatization that hurts workers ' welfare and does not take into account its social impact , " they said . The opposition party [15 15 Radical Cause 15] controls all of the unionized workers at the [16 16 CVG 16] heavy industry complex and has systematically opposed all government legislation in congress . -- Omar Lugo , [17 17 Caracas 17] newsroom , 582 834405

1: 32374 Venezuela
2: 31728475 Corporación_Venezolana_de_Guayana
3: 55289 Caracas
4: 32374 Venezuela
5: 31728475 Corporación_Venezolana_de_Guayana
6: 31728475 Corporación_Venezolana_de_Guayana
7: 31728475 Corporación_Venezolana_de_Guayana
8: 22013617 SIDOR
9: 2660398 Radical_Cause
10: 22013617 SIDOR
11: 31728475 Corporación_Venezolana_de_Guayana
12: 31728475 Corporación_Venezolana_de_Guayana
13: 31728475 Corporación_Venezolana_de_Guayana

14: 31728475 Corporación_Venezolana_de_Guayana
15: 2660398 Radical_Cause
16: 31728475 Corporación_Venezolana_de_Guayana
17: 55289 Caracas
1: 32374 Venezuela
[pem_pos=0, final_scores=0.56516355,
final_local_scores=0.632627, pem=-0.21226841,
similarity=24.122913]
2: 31728475 Corporación_Venezolana_de_Guayana
[pem_pos=1, final_scores=0.66339326,
final_local_scores=0.6351641, pem=-0.49405056,
similarity=25.730999]
3: 55289 Caracas
[pem_pos=0, final_scores=0.40593618,
final_local_scores=0.4554276, pem=0.0, similarity=17.68058]
4: 32374 Venezuela
[pem_pos=0, final_scores=0.5708712,
final_local_scores=0.6388618, pem=-0.21226841,
similarity=24.308971]
5: 31728475 Corporación_Venezolana_de_Guayana
[pem_pos=0, final_scores=0.5283628,
final_local_scores=0.48766387, pem=0.0, similarity=18.642576]
6: 31728475 Corporación_Venezolana_de_Guayana
[pem_pos=1, final_scores=0.5258693,
final_local_scores=0.48494005, pem=-0.49405056,
similarity=21.248014]
7: 31728475 Corporación_Venezolana_de_Guayana
[pem_pos=1, final_scores=0.5738523,
final_local_scores=0.5373542, pem=-0.49405056,
similarity=22.812157]
8: 22013617 SIDOR
[pem_pos=0, final_scores=0.4837759,
final_local_scores=0.58225113, pem=0.0, similarity=21.465248]
9: 2660398 Radical_Cause
[pem_pos=0, final_scores=0.4944157,
final_local_scores=0.5852214, pem=0.0, similarity=21.553886]
10: 22013617 SIDOR
[pem_pos=0, final_scores=0.4198373,
final_local_scores=0.51240796, pem=0.0, similarity=19.38099]
11: 31728475 Corporación_Venezolana_de_Guayana
[pem_pos=1, final_scores=0.6360517,
final_local_scores=0.6052976, pem=-0.49405056,
similarity=24.839725]
12: 31728475 Corporación_Venezolana_de_Guayana
[pem_pos=1, final_scores=0.60323805,
final_local_scores=0.56945366, pem=-0.49405056,
similarity=23.77007]
13: 31728475 Corporación_Venezolana_de_Guayana
[pem_pos=1, final_scores=0.6831302,
final_local_scores=0.6567237, pem=-0.49405056,
similarity=26.37438]
14: 31728475 Corporación_Venezolana_de_Guayana

```
[pem_pos=1, final_scores=0.66923803,
final_local_scores=0.64154863, pem=-0.49405056,
similarity=25.921528]
15: 2660398 Radical_Cause
    [pem_pos=0, final_scores=0.41787958,
final_local_scores=0.5016174, pem=0.0, similarity=19.058975]
16: 31728475 Corporación_Venezolana_de_Guayana
    [pem_pos=1, final_scores=0.618614,
final_local_scores=0.5862496, pem=-0.49405056,
similarity=24.271294]
17: 55289 Caracas
    [pem_pos=0, final_scores=0.46755838,
final_local_scores=0.5227405, pem=0.0, similarity=19.689331]
```

Sample: b'571_Nicaraguan'

[1 1 Nicaraguan 1] drunks fear " lovebite " bandit . [2 2 MA
NAGUA 2] , [3 3 Nicaragua 3] 1996-08-27 Heavy drinkers in a [4
4 Nicaraguan 4] city were searching for someone who has
covered them in " lovebites " while they were passed out in a
drunken stupor , a local newspaper reported on Tuesday . The
dreaded " chupabolos " -- " drunksucker " -- preys on men who
have passed out in the streets of [5 5 Matagalpa 5] , 80
miles (130 kms) north of [6 6 Managua 6] , placing hickey-
like " lovebites " on various parts of their bodies , [7 7 El
Nuevo Diario 7] reported . Enraged drunks and street people
in this town known for its machismo have organised a so-far
unsuccessful search for the culprit who finds victims in the
dark streets surrounding a local market . The total number of
victims was still unknown . The first of the victims were two
vagrants who slept in an abandoned car in front of a local
bank , the newspaper said . In spite of the collective fear
gripping [8 8 Matagalpa 8] 's drinkers , local women
expressed little sympathy . " Its just desserts for all the '
bolos ' (drunkards) who sleep in the streets of our
beautiful town , " said a woman who worked in the local
market .

```
1: 21362 Nicaragua
2: 57042 Managua
3: 21362 Nicaragua
4: 21362 Nicaragua
5: 2226978 Matagalpa
6: 57042 Managua
7: 7443389 El_Nuevo_Diario
8: 2226978 Matagalpa
1: 21362 Nicaragua
    [pem_pos=0, final_scores=0.4934004,
final_local_scores=0.45017856, pem=0.0, similarity=17.52394]
2: 57042 Managua
    [pem_pos=0, final_scores=0.44153553,
final_local_scores=0.40721583, pem=0.0, similarity=16.241846]
3: 21362 Nicaragua
    [pem_pos=0, final_scores=0.5356031,
```

```
final_local_scores=0.49627858, pem=-0.24729352,
similarity=20.244476]
4: 21362 Nicaragua
    [pem_pos=0, final_scores=0.52379626,
final_local_scores=0.48338145, pem=0.0, similarity=18.51478]
5: 2226978 Matagalpa
    [pem_pos=0, final_scores=0.31939268,
final_local_scores=0.30757952, pem=0.0, similarity=13.268499]
6: 57042 Managua
    [pem_pos=0, final_scores=0.40323722,
final_local_scores=0.3653807, pem=0.0, similarity=14.993402]
7: 7443389 El_Nuevo_Diario
    [pem_pos=0, final_scores=0.344343,
final_local_scores=0.46555686, pem=0.0, similarity=17.982859]
8: 2226978 Matagalpa
    [pem_pos=0, final_scores=0.37574887,
final_local_scores=0.3691401, pem=0.0, similarity=15.105589]
```

Sample: b'572_Brazil'

[1 1 Brazil 1] likely to turn Banespa federal bank - paper .
[2 2 SAO PAULO 2] 1996-08-27 [3 3 Brazil 3] is likely to turn
[4 4 Sao Paulo 4] state bank Banespa into a federal bank in a
prior step to privatization , according to unnamed government
sources , [5 5 0 Globo 5] daily said . The newspaper said the
[6 6 Central Bank 6] special administration of Banespa ends
in December 30 and after that the bank has to be liquidated
or turned into a federal bank since there are no conditions
to return Banespa to [7 7 Sao Paulo 7] state government . A [8
8 Central Bank 8] spokesman said he could not confirm or deny
the report . Banespa has been under central bank special
temporary administration since December 1994 . The central
bank management could be lifted if [9 9 Sao Paulo 9] state
decided to take part in the recent federal government 's plan
to restructure state banks . Under the plan , the federal
government would provide 100 percent of the financing needed
to restructure debt of state banks being privatized ,
liquidated or turned into development banks . It also offers
to refinance up to 50 percent of the debt held by state banks
whose governments decide to keep control of their banks .
Although the plan was designed under terms proposed by [10 10
Sao Paulo 10] state governor [11 11 Mario Covas 11] , he has
showed no interest in taking part in the plan because [12 12 S
ao Paulo 12] 's debt with Banespa has increased sharply , [13
13 0 Globo 13] said . [14 14 Sao Paulo 14] state 's debt is
now estimated at 19 billion reais . [15 15 0 Globo 15] also
said another delicate case to be solved involves private bank
Bamerindus . The newspaper said Bamerindus has sent to the [16
16 Central Bank 16] a proposal for restructuring combined
with a request for a 90-day credit line , paying four percent
a year plus the Basic Interest Rate of the [17 17 Central Bank
17] (TBC) . [18 18 0 Globo 18] also said the loan would
give Bamerindus time to sell assets . Bamerindus , [19 19 Bra

`zil 19] 's fourth-largest private bank , has been facing liquidity troubles . Bamerindus declined to comment on negotiations being held with the [20 20 Central Bank 20] . -- Fatima Cristina , [21 21 São Paulo 21] newsroom , 55-11-2324411`

`1: 3383 Brazil`

`2: 390875 São_Paulo`

`3: 3383 Brazil`

`4: 229422 São_Paulo_(state)`

`5: 2093729 O_Globo`

`6: 917127 Central_Bank_of_Brazil`

`7: 229422 São_Paulo_(state)`

`8: 917127 Central_Bank_of_Brazil`

`9: 229422 São_Paulo_(state)`

`10: 229422 São_Paulo_(state)`

`11: 3386439 Mário_Covas`

`12: 229422 São_Paulo_(state)`

`13: 2093729 O_Globo`

`14: 229422 São_Paulo_(state)`

`15: 2093729 O_Globo`

`16: 917127 Central_Bank_of_Brazil`

`17: 917127 Central_Bank_of_Brazil`

`18: 2093729 O_Globo`

`19: 3383 Brazil`

`20: 917127 Central_Bank_of_Brazil`

`21: 390875 São_Paulo`

`1: 3383 Brazil`

`[pem_pos=0, final_scores=0.7024936, final_local_scores=0.71248424, pem=-0.21313342, similarity=26.510715]`

`2: 390875 São_Paulo`

`[pem_pos=0, final_scores=0.3906179, final_local_scores=0.35587114, pem=0.0, similarity=14.709618]`

`3: 3383 Brazil`

`[pem_pos=0, final_scores=0.68485993, final_local_scores=0.69322205, pem=-0.21313342, similarity=25.935894]`

`4: 229422 São_Paulo_(state)`

`[pem_pos=1, final_scores=0.40330803, final_local_scores=0.36534756, pem=-1.5484804, similarity=23.413292]`

`5: 2093729 O_Globo`

`[pem_pos=0, final_scores=0.5547717, final_local_scores=0.59581983, pem=0.0, similarity=21.870165]`

`6: 917127 Central_Bank_of_Brazil`

`[pem_pos=1, final_scores=0.47156543, final_local_scores=0.50996333, pem=-1.5394354, similarity=27.679724]`

`7: 229422 São_Paulo_(state)`

`[pem_pos=1, final_scores=0.46922392, final_local_scores=0.4373507, pem=-1.5484804, similarity=25.562008]`

```
8: 917127 Central_Bank_of_Brazil
    [pem_pos=1, final_scores=0.387114,
final_local_scores=0.417713, pem=-1.5394354,
similarity=24.926792]
9: 229422 São_Paulo_(state)
    [pem_pos=1, final_scores=0.445948,
final_local_scores=0.41192526, pem=-1.5484804,
similarity=24.80326]
10: 229422 São_Paulo_(state)
    [pem_pos=1, final_scores=0.4465013,
final_local_scores=0.41252965, pem=-1.5484804,
similarity=24.821299]
11: 3386439 Mário_Covas
    [pem_pos=0, final_scores=0.55213445,
final_local_scores=0.6519886, pem=0.0, similarity=23.546354]
12: 229422 São_Paulo_(state)
    [pem_pos=1, final_scores=0.37585813,
final_local_scores=0.33536276, pem=-1.5484804,
similarity=22.518484]
13: 2093729 0_Globo
    [pem_pos=0, final_scores=0.5204802,
final_local_scores=0.5583615, pem=0.0, similarity=20.752333]
14: 229422 São_Paulo_(state)
    [pem_pos=1, final_scores=0.43979758,
final_local_scores=0.40520686, pem=-1.5484804,
similarity=24.602772]
15: 2093729 0_Globo
    [pem_pos=0, final_scores=0.5301418,
final_local_scores=0.5689153, pem=0.0, similarity=21.06728]
16: 917127 Central_Bank_of_Brazil
    [pem_pos=1, final_scores=0.4069023,
final_local_scores=0.43932873, pem=-1.5394354,
similarity=25.571846]
17: 917127 Central_Bank_of_Brazil
    [pem_pos=1, final_scores=0.43898886,
final_local_scores=0.4743784, pem=-1.5394354,
similarity=26.6178]
18: 2093729 0_Globo
    [pem_pos=0, final_scores=0.5232691,
final_local_scores=0.561408, pem=0.0, similarity=20.843246]
19: 3383 Brazil
    [pem_pos=0, final_scores=0.7063387,
final_local_scores=0.71668434, pem=-0.21313342,
similarity=26.636055]
20: 917127 Central_Bank_of_Brazil
    [pem_pos=1, final_scores=0.37053853,
final_local_scores=0.39960688, pem=-1.5394354,
similarity=24.386469]
21: 390875 São_Paulo
    [pem_pos=0, final_scores=0.32361728,
final_local_scores=0.28268316, pem=0.0, similarity=12.525542]
```

Sample: b'573_Czech'

[1 1 Czech Republic 1] 's [2 2 Vaclav Havel 2] to tour [3 3 Brazil 3] in September . [4 4 BRASILIA 4] 1996-08-27 [5 5 Czech Republic 5] President [6 6 Vaclav Havel 6] is scheduled to make an official visit to [7 7 Brazil 7] Sept . 15-21 , [8 8 Brazil 8] 's Foreign Relations Ministry said on Tuesday . [9 9 Havel 9] is due to meet with his [10 10 Brazilian 10] counterpart [11 11 Fernando Henrique Cardoso 11] in the capital [12 12 Brasilia 12] and will visit the cities of Manaus , [13 13 Sao Paulo 13] and [14 14 Rio de Janeiro 14] . Also due to visit [15 15 Brazil 15] in September are [16 16 South Korean 16] President [17 17 Kim Young Sam 17] and [18 18 German 18] Chancellor [19 19 Helmut Kohl 19] .

1: 5321 Czech_Republic
2: 63299 Václav_Havel
3: 3383 Brazil
4: 4752 Brasília
5: 5321 Czech_Republic
6: 63299 Václav_Havel
7: 3383 Brazil
8: 3383 Brazil
9: 63299 Václav_Havel
10: 3383 Brazil
11: 82491 Fernando_Henrique_Cardoso
12: 4752 Brasília
13: 390875 São_Paulo
14: 25936 Rio_de_Janeiro
15: 3383 Brazil
16: 27019 South_Korea
17: 35991378 Kim_Young-sam
18: 11867 Germany
19: 41896 Helmut_Kohl
1: 5321 Czech_Republic
[pem_pos=0, final_scores=0.5968836,
final_local_scores=0.5971885, pem=-0.27794555,
similarity=23.42252]
2: 63299 Václav_Havel
[pem_pos=0, final_scores=0.7518925,
final_local_scores=0.8059434, pem=0.0, similarity=28.140673]
3: 3383 Brazil
[pem_pos=0, final_scores=0.7047505,
final_local_scores=0.70732754, pem=-0.21313342,
similarity=26.35683]
4: 4752 Brasília
[pem_pos=0, final_scores=0.5135917,
final_local_scores=0.5069653, pem=0.0, similarity=19.218569]
5: 5321 Czech_Republic
[pem_pos=0, final_scores=0.57127744,
final_local_scores=0.56921756, pem=-0.27794555,
similarity=22.587812]
6: 63299 Václav_Havel
[pem_pos=0, final_scores=0.7182557,

```
final_local_scores=0.7573475, pem=0.0, similarity=26.690474]
7: 3383 Brazil
    [pem_pos=0, final_scores=0.728561,
final_local_scores=0.7333369, pem=-0.21313342,
similarity=27.133001]
8: 3383 Brazil
    [pem_pos=0, final_scores=0.7760597,
final_local_scores=0.785222, pem=-0.21313342,
similarity=28.681356]
9: 63299 Václav_Havel
    [pem_pos=0, final_scores=0.68825704,
final_local_scores=0.7364313, pem=0.0, similarity=26.066292]
10: 3383 Brazil
    [pem_pos=0, final_scores=0.6385544,
final_local_scores=0.63501835, pem=-0.2733083,
similarity=24.52622]
11: 82491 Fernando_Henrique_Cardoso
    [pem_pos=0, final_scores=0.5622435,
final_local_scores=0.61262393, pem=0.0, similarity=22.371634]
12: 4752 Brasília
    [pem_pos=0, final_scores=0.33748186,
final_local_scores=0.314592, pem=0.0, similarity=13.477766]
13: 390875 São_Paulo
    [pem_pos=0, final_scores=0.34735388,
final_local_scores=0.3205256, pem=0.0, similarity=13.654836]
14: 25936 Rio_de_Janeiro
    [pem_pos=0, final_scores=0.4156266,
final_local_scores=0.3884428, pem=-0.1260924,
similarity=16.36733]
15: 3383 Brazil
    [pem_pos=0, final_scores=0.6383118,
final_local_scores=0.63475347, pem=-0.21313342,
similarity=24.191074]
16: 27019 South_Korea
    [pem_pos=0, final_scores=0.59372175,
final_local_scores=0.6191189, pem=0.0, similarity=22.565456]
17: 35991378 Kim_Young-sam
    [pem_pos=0, final_scores=0.38729537,
final_local_scores=0.48372233, pem=0.0, similarity=18.524952]
18: 11867 Germany
    [pem_pos=1, final_scores=0.6946158,
final_local_scores=0.70530456, pem=-1.0465823,
similarity=30.828888]
19: 41896 Helmut_Kohl
    [pem_pos=0, final_scores=0.73899263,
final_local_scores=0.8342671, pem=0.0, similarity=28.985907]
```

Sample: b'574_Formal'

Former [1 1 Argentine 1] benevolent dictator [2 2 Alejandro_Lanusse 2] dies . [3 3 BUENOS_AIRES 3] 1996-08-26 [4 4 Alejandro_Lanusse 4] , the former dictator who ruled [5 5 Argentina 5] for two years , died at age 78 on Monday . [6 6 Lanusse 6]

died after being brought to a hospital a week ago following a fall at home that resulted in a blood clot in the brain . He was operated on earlier in the week but failed to recover from surgery . The former dictator , who ruled from 1971 to 1973 , was best known for allowing [7 7 Juan Domingo Peron 7] , [8 8 Argentina 8] 's famed populist leader , to return to [9 9 Argentina 9] after 17 years of forced exile . [10 10 Lanusse 10] took over the leadership of the country after five years of dictatorship . But unlike his two predecessors , Juan Carlos Onganía and [11 11 Marcelo Levingston 11] , who ruled [12 12 Argentina 12] with an iron hand , [13 13 Lanusse 13] steered the country toward democracy . That resulted in general elections in March 1973 when the [14 14 Peronists 14] led by Hector Campora and [15 15 Vicente Solano Lima 15] returned to power . [16 16 Lanusse 16] was a candidate in the election but failed to defeat his old adversaries and never returned to public office . He was imprisoned for four years in 1951 for taking part in a coup attempt to overthrow [17 17 Peron 17] led by General Benjamin Menendez . [18 18 Lanusse 18] 's rule saw the gradual rise of left-wing activism which culminated in another period of brutal [19 19 Argentine 19] dictatorship from 1976 to 1983 , during which the military launched its " dirty war " that resulted in 10,000 missing people . In his autobiography published in 1990 , [20 20 Lanusse 20] described himself as a military man with " democratic ideas . " He was born in [21 21 Buenos Aires 21] in 1918 and married Ileana Bell with whom he had nine children . He entered the Military College in 1935 .

1: 18951905 Argentina
2: 339667 Alejandro_Agustín_Lanusse
3: 255919 Buenos_Aires
4: 339667 Alejandro_Agustín_Lanusse
5: 18951905 Argentina
6: 339667 Alejandro_Agustín_Lanusse
7: 23486968 Juan_Perón
8: 18951905 Argentina
9: 18951905 Argentina
10: 339667 Alejandro_Agustín_Lanusse
11: 339668 Roberto_M._Levingston
12: 18951905 Argentina
13: 339667 Alejandro_Agustín_Lanusse
14: 242379 Peronism
15: 1671729 Vicente_Solano_Lima
16: 339667 Alejandro_Agustín_Lanusse
17: 23486968 Juan_Perón
18: 339667 Alejandro_Agustín_Lanusse
19: 18951905 Argentina
20: 339667 Alejandro_Agustín_Lanusse
21: 255919 Buenos_Aires
1: 18951905 Argentina
[pem_pos=0, final_scores=0.5434299,
final_local_scores=0.5834933, pem=-0.18885723,

```
similarity=22.529352]
2: 339667 Alejandro_Agustín_Lanusse
    [pem_pos=0, final_scores=0.58474237,
final_local_scores=0.54724723, pem=0.0, similarity=20.420662]
3: 255919 Buenos_Aires
    [pem_pos=0, final_scores=0.44754952,
final_local_scores=0.4608652, pem=-0.12597719,
similarity=18.527933]
4: 339667 Alejandro_Agustín_Lanusse
    [pem_pos=0, final_scores=0.5287064,
final_local_scores=0.48603636, pem=0.0, similarity=18.594007]
5: 18951905 Argentina
    [pem_pos=0, final_scores=0.5171564,
final_local_scores=0.5547935, pem=-0.30003807,
similarity=22.27751]
6: 339667 Alejandro_Agustín_Lanusse
    [pem_pos=0, final_scores=0.7565473,
final_local_scores=0.73810595, pem=0.0, similarity=26.116266]
7: 23486968 Juan_Perón
    [pem_pos=0, final_scores=0.45959133,
final_local_scores=0.43702227, pem=0.0, similarity=17.131329]
8: 18951905 Argentina
    [pem_pos=0, final_scores=0.5636272,
final_local_scores=0.60555583, pem=-0.30003807,
similarity=23.79236]
9: 18951905 Argentina
    [pem_pos=0, final_scores=0.5830448,
final_local_scores=0.62676656, pem=-0.30003807,
similarity=24.425331]
10: 339667 Alejandro_Agustín_Lanusse
    [pem_pos=0, final_scores=0.7129815,
final_local_scores=0.69051695, pem=0.0, similarity=24.696115]
11: 339668 Roberto_M._Levingston
    [pem_pos=0, final_scores=0.48674464,
final_local_scores=0.56845003, pem=0.0, similarity=21.053396]
12: 18951905 Argentina
    [pem_pos=0, final_scores=0.5106682,
final_local_scores=0.5477061, pem=-0.30003807,
similarity=22.06601]
13: 339667 Alejandro_Agustín_Lanusse
    [pem_pos=0, final_scores=0.6764955,
final_local_scores=0.6506615, pem=0.0, similarity=23.50675]
14: 242379 Peronism
    [pem_pos=0, final_scores=0.24509083,
final_local_scores=0.21957693, pem=0.0, similarity=10.642326]
15: 1671729 Vicente_Solano_Lima
    [pem_pos=0, final_scores=0.48953068,
final_local_scores=0.54087085, pem=0.0, similarity=20.230377]
16: 339667 Alejandro_Agustín_Lanusse
    [pem_pos=0, final_scores=0.6635454,
final_local_scores=0.6365155, pem=0.0, similarity=23.084604]
17: 23486968 Juan_Perón
```

```
[pem_pos=0, final_scores=0.44783664,
final_local_scores=0.42944717, pem=0.0, similarity=16.905273]
18: 339667 Alejandro_Agustín_Lanusse
    [pem_pos=0, final_scores=0.69922477,
final_local_scores=0.6754898, pem=0.0, similarity=24.247675]
19: 18951905 Argentina
    [pem_pos=0, final_scores=0.49741632,
final_local_scores=0.5332304, pem=-0.18885723,
similarity=21.029408]
20: 339667 Alejandro_Agustín_Lanusse
    [pem_pos=0, final_scores=0.6489039,
final_local_scores=0.6205219, pem=0.0, similarity=22.607325]
21: 255919 Buenos_Aires
    [pem_pos=0, final_scores=0.39942604,
final_local_scores=0.4082976, pem=-0.12597719,
similarity=16.959211]
```

Sample: b'575_Ten'

Ten missing in north [1 1 China 1] ship collision . [2 2 BEI JING 2] 1996-08-27 Ten people were missing after a fishing boat collided with a passenger liner and sank off [3 3 China 3] 's northeastern province of [4 4 Liaoning 4] , state radio said on Tuesday . The fishing boat sank and its entire crew was missing after a collision with the " Tiantan " liner off the port of [5 5 Dalian 5] early on Monday , the report said . It said the liner was heading to [6 6 Dalian 6] from the northern port of [7 7 Tianjin 7] , it said . [8 8 Dalian 8] port officials , contacted by telephone , confirmed the collision but gave no further details .

```
1: 5405 China
2: 18603746 Beijing
3: 5405 China
4: 155516 Liaoning
5: 250109 Dalian
6: 250109 Dalian
7: 50821 Tianjin
8: 250109 Dalian
1: 5405 China
    [pem_pos=0, final_scores=0.526061,
final_local_scores=0.49920106, pem=-0.15669827,
similarity=19.839018]
2: 18603746 Beijing
    [pem_pos=0, final_scores=0.526182,
final_local_scores=0.51849836, pem=-0.042161036,
similarity=19.792017]
3: 5405 China
    [pem_pos=0, final_scores=0.48320562,
final_local_scores=0.45238805, pem=-0.15669827,
similarity=18.442024]
4: 155516 Liaoning
    [pem_pos=0, final_scores=0.38819766,
final_local_scores=0.3497523, pem=0.0, similarity=14.5270195]
```

```
5: 250109 Dalian
    [pem_pos=0, final_scores=0.4383505,
final_local_scores=0.3810643, pem=0.0, similarity=15.461432]
6: 250109 Dalian
    [pem_pos=0, final_scores=0.47982132,
final_local_scores=0.4263649, pem=0.0, similarity=16.813292]
7: 50821 Tianjin
    [pem_pos=0, final_scores=0.43660277,
final_local_scores=0.38564795, pem=0.0, similarity=15.598217]
8: 250109 Dalian
    [pem_pos=0, final_scores=0.37997526,
final_local_scores=0.31729823, pem=0.0, similarity=13.558525]
```

Sample: b'576_Matahari'

Matahari revises down 1996 net target . [1 1 JAKARTA 1]
1996-08-27 [2 2 Indonesian 2] department store operator PT Matahari Putra Prima said on Tuesday that it had revised down its 1996 net profit target . Matahari 's finance director , Hanifah Komala , said they revised down net profit for 1996 to 46 billion from its original target of 50 billion rupiah . " We have to revised down our target due to weak sales performance in the third quarter , " said Komala . He also said the company expect to record a 83 billion rupiah of net profit in 1997 . -- [3 3 Jakarta 3] newsroom +6221 384-6364

```
1: 16275 Jakarta
2: 14579 Indonesia
3: 16275 Jakarta
1: 16275 Jakarta
    [pem_pos=0, final_scores=0.4365644,
final_local_scores=0.36737746, pem=0.0, similarity=15.052989]
2: 14579 Indonesia
    [pem_pos=1, final_scores=0.48404104,
final_local_scores=0.47327995, pem=-0.74208045,
similarity=22.24888]
3: 16275 Jakarta
    [pem_pos=0, final_scores=0.5412846,
final_local_scores=0.48176843, pem=0.0, similarity=18.466644]
```

Sample: b'577_S.'

[1 1 S. Korea 1] asked to stop China-bound missionaries . [2 2 BEIJING 2] 1996-08-27 [3 3 Beijing 3] has called on [4 4 Seoul 4] to stop [5 5 South Korean 5] missionaries from travelling to [6 6 China 6] , a [7 7 South Korean 7] embassy spokesman said on Tuesday . The appeal was made on Sunday during talks between [8 8 South Korean 8] deputy Foreign Minister Lee Ki-choo and his [9 9 Chinese 9] counterpart [10 10 Tang Jiaxuan 10] , the spokesman said . It is not known why [11 11 China 11] raised the issue . Atheist [12 12 China 12] officially bans missionary activities but often turns a blind eye to religious activities of people nominally employed as foreign language teachers , particularly in remote areas that are unable to attract other candidates .

```
1: 27019 South_Korea
2: 18603746 Beijing
3: 5405 China
4: 19159283 Seoul
5: 27019 South_Korea
6: 5405 China
7: 27019 South_Korea
8: 27019 South_Korea
9: 5405 China
10: 1585055 Tang_Jiaxuan
11: 5405 China
12: 5405 China
1: 27019 South_Korea
    [pem_pos=0, final_scores=0.52501184,
final_local_scores=0.4807173, pem=0.0, similarity=18.435276]
2: 18603746 Beijing
    [pem_pos=0, final_scores=0.5161043,
final_local_scores=0.50166076, pem=-0.042161036,
similarity=19.289549]
3: 5405 China
    [pem_pos=7, final_scores=0.28598547,
final_local_scores=0.21334255, pem=-3.4768412,
similarity=29.363884]
4: 19159283 Seoul
    [pem_pos=0, final_scores=0.44258988,
final_local_scores=0.4148605, pem=-0.0378014,
similarity=16.675549]
5: 27019 South_Korea
    [pem_pos=0, final_scores=0.511527,
final_local_scores=0.46598715, pem=0.0, similarity=17.995699]
6: 5405 China
    [pem_pos=0, final_scores=0.4711209,
final_local_scores=0.41557485, pem=-0.15669827,
similarity=17.343445]
7: 27019 South_Korea
    [pem_pos=0, final_scores=0.59886545,
final_local_scores=0.5613912, pem=0.0, similarity=20.842745]
8: 27019 South_Korea
    [pem_pos=0, final_scores=0.58757514,
final_local_scores=0.5490582, pem=0.0, similarity=20.474705]
9: 5405 China
    [pem_pos=1, final_scores=0.6167193,
final_local_scores=0.574619, pem=-1.4196125,
similarity=28.957565]
10: 1585055 Tang_Jiaxuan
    [pem_pos=0, final_scores=0.48917288,
final_local_scores=0.63906896, pem=0.0, similarity=23.160805]
11: 5405 China
    [pem_pos=0, final_scores=0.5528937,
final_local_scores=0.50489926, pem=-0.15669827,
similarity=20.009064]
12: 5405 China
```

```
[pem_pos=0, final_scores=0.54081947,
final_local_scores=0.49170995, pem=-0.15669827,
similarity=19.615469]
```

Sample: b'578_Hong'

[1 1 Hong Kong 1] nabs blind 10-year-old illegal immigrant .
[2 2 HONG KONG 2] 1996-08-27 A blind 10-year-old boy from [3 3]
China 3] sneaked over the border into [4 4 Hong Kong 4] and
was arrested as an illegal immigrant , [5 5 Hong Kong 5]
police said on Tuesday . He was caught by police trying to
force his way into a home in the rural [6 6 New Territories 6]
, a police spokesman said . " The boy came from [7 7 China 7]
's eastern province of [8 8 Jiangsu 8] . He was spotted by a
passerby trying to climb into an apartment in the early hours
of Monday morning , " the spokesman said . No decision has
yet been made on how to deal with the boy . [9 9 Hong Kong 9]
police regularly catch hundreds of illegal immigrants and
people who have overstayed their visas from mainland [10 10 Ch
ina 10] and send them back . [11 11 Hong Kong 11] , a [12 12 B
ritish 12] colony , reverts to [13 13 Chinese 13] control
next year but will remain sealed off from the mainland except
to a tiny trickle of legal immigrants and people with special
visit permits .

1: 13404 Hong_Kong

2: 13404 Hong_Kong

3: 5405 China

4: 13404 Hong_Kong

5: 13404 Hong_Kong

6: 230930 New_Territories

7: 5405 China

8: 166412 Jiangsu

9: 13404 Hong_Kong

10: 5405 China

11: 13404 Hong_Kong

12: 31717 United_Kingdom

13: 5405 China

1: 13404 Hong_Kong

```
[pem_pos=0, final_scores=0.5451003,
final_local_scores=0.48421896, pem=-0.22402541,
similarity=19.758059]
```

2: 13404 Hong_Kong

```
[pem_pos=0, final_scores=0.5987258,
final_local_scores=0.5427967, pem=-0.22402541,
similarity=21.506134]
```

3: 5405 China

```
[pem_pos=0, final_scores=0.5749001,
final_local_scores=0.53656715, pem=-0.15669827,
similarity=20.954098]
```

4: 13404 Hong_Kong

```
[pem_pos=0, final_scores=0.50780684,
final_local_scores=0.44348156, pem=-0.22402541,
similarity=18.542374]
```

```
5: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.51335096,
final_local_scores=0.44953763, pem=-0.22402541,
similarity=18.723099]
6: 230930 New_Territories
    [pem_pos=0, final_scores=0.506422,
final_local_scores=0.52986026, pem=0.0, similarity=19.9018]
7: 5405 China
    [pem_pos=0, final_scores=0.5701571,
final_local_scores=0.5313862, pem=-0.15669827,
similarity=20.799486]
8: 166412 Jiangsu
    [pem_pos=0, final_scores=0.42360413,
final_local_scores=0.4381739, pem=0.0, similarity=17.165695]
9: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5732675,
final_local_scores=0.5149874, pem=-0.22402541,
similarity=20.67625]
10: 5405 China
    [pem_pos=0, final_scores=0.49774426,
final_local_scores=0.45228618, pem=-0.15669827,
similarity=18.438984]
11: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5536069,
final_local_scores=0.49351114, pem=-0.22402541,
similarity=20.035355]
12: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.561987,
final_local_scores=0.56710404, pem=-0.48936236,
similarity=23.674458]
13: 5405 China
    [pem_pos=1, final_scores=0.74764436,
final_local_scores=0.725264, pem=-1.4196125,
similarity=33.453114]
```

Sample: b'579_Bosnian'

Bosnian premier in Turkey for one day visit . [1 1 ANKARA 1]
1996-08-27 Bosnian Prime Minister [2 2 Hasan Muratovic 2]
arrived in [3 3 Ankara 3] on Tuesday for an official visit
where he is due to discuss Turkey 's aid to the former [4 4 Yu
goslav 4] republic . The premier , who is due to meet his
Turkish counterpart [5 5 Necmettin Erbakan 5] on Tuesday ,
will also be discussing the postponed Bosnian elections , a
foreign ministry official said . A small number of Bosnians
had also begun to vote in Turkey . Muratovic is also due to
meet with President [6 6 Suleyman Demirel 6] , Foreign
Minister [7 7 Tansu Ciller 7] and Turkish businessman , the
ministry official said . He will leave on Thursday . A [8 8 U.
S. 8] diplomat in charge of elections in Bosnia announced
earlier that municipal polls due on September 14 with other
Bosnian elections would be put off because of irregularities
by the [9 9 Serbs 9] in registering voters . He said no new

date had been set yet . " Turkish people are watching closely the developments in Bosnia . We have seen elections as a step in the normalisation process , " the foreign ministry official said .

1: 802 Ankara
2: 25887840 Hasan_Muratović
3: 802 Ankara
4: 34244 Yugoslavia
5: 1328699 Necmettin_Erbakan
6: 166093 Süleyman_Demirel
7: 753267 Tansu_Çiller
8: 3434750 United_States
9: 29265 Serbia
1: 802 Ankara
 [pem_pos=0, final_scores=0.433504,
final_local_scores=0.40246934, pem=0.0, similarity=16.1002]
2: 25887840 Hasan_Muratović
 [pem_pos=0, final_scores=0.42338198,
final_local_scores=0.57615757, pem=0.0, similarity=21.283403]
3: 802 Ankara
 [pem_pos=0, final_scores=0.35866934,
final_local_scores=0.3207239, pem=0.0, similarity=13.660753]
4: 34244 Yugoslavia
 [pem_pos=0, final_scores=0.38646752,
final_local_scores=0.36437374, pem=-0.7596219,
similarity=19.094294]
5: 1328699 Necmettin_Erbakan
 [pem_pos=0, final_scores=0.56606776,
final_local_scores=0.63430107, pem=0.0, similarity=23.018522]
6: 166093 Süleyman_Demirel
 [pem_pos=0, final_scores=0.6053534,
final_local_scores=0.62688446, pem=0.0, similarity=22.797195]
7: 753267 Tansu_Çiller
 [pem_pos=0, final_scores=0.7803559,
final_local_scores=0.85780907, pem=0.0, similarity=29.688446]
8: 3434750 United_States
 [pem_pos=0, final_scores=0.4038754,
final_local_scores=0.44652432, pem=-0.14075686,
similarity=18.180346]
9: 29265 Serbia
 [pem_pos=4, final_scores=0.5019814,
final_local_scores=0.5214644, pem=-2.3277044,
similarity=32.30967]

Sample: b'580_New'

New [1 1 U.N. 1] relief coordinator arrives in [2 2 Iraq
2] . Leon Barkho [3 3 BAGHDAD 3] 1996-08-27 A new [4 4 U.N. 4]
relief coordinator has arrived in [5 5 Baghdad 5] to take up
the task of organising humanitarian goods distribution and to
face [6 6 Iraq 6] 's continuing opposition over the number of
international monitors to be involved . [7 7 U.N. 7] and
diplomatic sources said on Tuesday that Secretary-General [8

8 Boutros Boutros-Ghali 8] had appointed [9 9 Italian 9] Gualtiero Fulcheri and sent him to [10 10 Iraq 10] last week to replace [11 11 Moroccan 11] Mohamed Zejjari . One diplomat said [12 12 Iraq 12] and [13 13 U.N. 13] were still in disagreement on how many international observers would be required to ascertain the equitable distribution of humanitarian supplies that will be procured under [14 14 Baghdad 14] 's oil deal with U.N . " The [15 15 United Nations 15] would like to employ hundreds of foreign monitors . [16 16 Baghdad 16] says it can only accept a few dozens , " said the diplomat . [17 17 Baghdad 17] holds that the Iraq-U.N. memorandum of understanding on partial oil sales signed last June does not specify how many foreign observers should be stationed in [18 18 Iraq 18] . " Observation of food supplies and their distribution are still a major issue and seems the two sides have not yet filled the gap separating them , " said another diplomat . [19 19 Iraq 19] 's partial oil sales pact with [20 20 U.N. 20] , allowing crude exports worth \$ 2 billion every six months , gives [21 21 U.N. 21] the right to supervise the purchase and distribution of food supplies in the country . The deal is a humanitarian exception to the [22 22 U.N. 22] sanctions imposed on [23 23 Iraq 23] for invading [24 24 Kuwait 24] in 1990 which include a ban on its oil exports . Fulcheri declined comment on the differences between the [25 25 U.N. 25] and [26 26 Iraq 26] , saying only : " There are several different things which still need to be done . " Fulcheri started his [27 27 U.N. 27] career in 1960 and has long experience in [28 28 U.N. 28] emergency relief in [29 29 Congo 29] , [30 30 Angola 30] , [31 31 Sudan 31] and [32 32 Somalia 32] .

1: 31769 United_Nations
2: 7515928 Iraq
3: 4492 Baghdad
4: 31769 United_Nations
5: 4492 Baghdad
6: 7515928 Iraq
7: 31769 United_Nations
8: 211193 Boutros_Boutros-Ghali
9: 14532 Italy
10: 7515928 Iraq
11: 19291 Morocco
12: 7515928 Iraq
13: 31769 United_Nations
14: 4492 Baghdad
15: 31769 United_Nations
16: 4492 Baghdad
17: 4492 Baghdad
18: 7515928 Iraq
19: 7515928 Iraq
20: 31769 United_Nations
21: 31769 United_Nations
22: 31769 United_Nations

```
23: 7515928 Iraq
24: 7515890 Kuwait
25: 31769 United_Nations
26: 7515928 Iraq
27: 31769 United_Nations
28: 31769 United_Nations
29: 19599929 Republic_of_the_Congo
30: 701 Angola
31: 27421 Sudan
32: 27358 Somalia
1: 31769 United_Nations
    [pem_pos=0, final_scores=0.51581234,
final_local_scores=0.48110002, pem=0.0, similarity=18.446697]
2: 7515928 Iraq
    [pem_pos=0, final_scores=0.5872124,
final_local_scores=0.5434253, pem=-0.25379938,
similarity=21.68681]
3: 4492 Baghdad
    [pem_pos=0, final_scores=0.4357735,
final_local_scores=0.41385335, pem=0.0, similarity=16.439922]
4: 31769 United_Nations
    [pem_pos=0, final_scores=0.48847222,
final_local_scores=0.45123512, pem=0.0, similarity=17.55547]
5: 4492 Baghdad
    [pem_pos=0, final_scores=0.3874324,
final_local_scores=0.361048, pem=0.0, similarity=14.864106]
6: 7515928 Iraq
    [pem_pos=0, final_scores=0.6420698,
final_local_scores=0.6033487, pem=-0.25379938,
similarity=23.475042]
7: 31769 United_Nations
    [pem_pos=0, final_scores=0.5068255,
final_local_scores=0.47128326, pem=0.0, similarity=18.153746]
8: 211193 Boutros_Boutros-Ghali
    [pem_pos=0, final_scores=0.7014943,
final_local_scores=0.75488317, pem=0.0, similarity=26.616932]
9: 14532 Italy
    [pem_pos=0, final_scores=0.6535782,
final_local_scores=0.7097425, pem=-0.75948656,
similarity=29.40005]
10: 7515928 Iraq
    [pem_pos=0, final_scores=0.55582416,
final_local_scores=0.5091384, pem=-0.25379938,
similarity=20.66362]
11: 19291 Morocco
    [pem_pos=0, final_scores=0.6962696,
final_local_scores=0.71002513, pem=-0.2674848,
similarity=26.732903]
12: 7515928 Iraq
    [pem_pos=0, final_scores=0.65235007,
final_local_scores=0.6145783, pem=-0.25379938,
similarity=23.810156]
```

```
13: 31769 United_Nations
    [pem_pos=0, final_scores=0.5062475,
final_local_scores=0.47065192, pem=0.0, similarity=18.134905]
14: 4492 Baghdad
    [pem_pos=0, final_scores=0.3233869,
final_local_scores=0.29108807, pem=0.0, similarity=12.776361]
15: 31769 United_Nations
    [pem_pos=0, final_scores=0.46295995,
final_local_scores=0.42336684, pem=0.0, similarity=16.723824]
16: 4492 Baghdad
    [pem_pos=0, final_scores=0.3072374,
final_local_scores=0.2734472, pem=0.0, similarity=12.249923]
17: 4492 Baghdad
    [pem_pos=0, final_scores=0.3149994,
final_local_scores=0.28192604, pem=0.0, similarity=12.502948]
18: 7515928 Iraq
    [pem_pos=0, final_scores=0.57695204,
final_local_scores=0.53221744, pem=-0.25379938,
similarity=21.352345]
19: 7515928 Iraq
    [pem_pos=0, final_scores=0.56339914,
final_local_scores=0.51741296, pem=-0.25379938,
similarity=20.91055]
20: 31769 United_Nations
    [pem_pos=0, final_scores=0.47302443,
final_local_scores=0.43436074, pem=0.0, similarity=17.051905]
21: 31769 United_Nations
    [pem_pos=0, final_scores=0.5222605,
final_local_scores=0.48814368, pem=0.0, similarity=18.656895]
22: 31769 United_Nations
    [pem_pos=0, final_scores=0.5506564,
final_local_scores=0.5191619, pem=0.0, similarity=19.582539]
23: 7515928 Iraq
    [pem_pos=0, final_scores=0.59696895,
final_local_scores=0.5540829, pem=-0.25379938,
similarity=22.004852]
24: 7515890 Kuwait
    [pem_pos=0, final_scores=0.46369213,
final_local_scores=0.42985928, pem=-0.3372397,
similarity=18.751534]
25: 31769 United_Nations
    [pem_pos=0, final_scores=0.5557663,
final_local_scores=0.5247436, pem=0.0, similarity=19.74911]
26: 7515928 Iraq
    [pem_pos=0, final_scores=0.62333536,
final_local_scores=0.58288413, pem=-0.25379938,
similarity=22.864338]
27: 31769 United_Nations
    [pem_pos=0, final_scores=0.5363119,
final_local_scores=0.5034927, pem=0.0, similarity=19.11494]
28: 31769 United_Nations
    [pem_pos=0, final_scores=0.5322738,
```

```
final_local_scores=0.49908167, pem=0.0, similarity=18.983305]
29: 19599929 Republic_of_the_Congo
    [pem_pos=1, final_scores=0.09996624,
final_local_scores=0.11330736, pem=-1.6516278,
similarity=16.452839]
30: 701 Angola
    [pem_pos=0, final_scores=0.5481298,
final_local_scores=0.57248026, pem=-0.39249876,
similarity=23.308136]
31: 27421 Sudan
    [pem_pos=0, final_scores=0.5234601,
final_local_scores=0.49021566, pem=-0.2848764,
similarity=20.26793]
32: 27358 Somalia
    [pem_pos=0, final_scores=0.58057195,
final_local_scores=0.5677776, pem=-0.28429922,
similarity=22.579391]
```

Sample: b'581_Kansas'

```
[1 1 Kansas 1] feedlot cattle market quiet , no sales - [2 2
USDA 2] . [3 3 DODGE CITY 3] 1996-08-27 Trade was quiet ,
with no sales slaughter steers or heifers confirmed . Inquiry
and demand very light . Sales confirmed week to date on 4,200
head , mostly previously contracted or formulated cattle .
Confirmed - none . -- [4 4 Chicago 4] newsdesk 312 408 8720--
1: 16716 Kansas
2: 70896 United_States_Department_of_Agriculture
3: 114560 Dodge_City,_Kansas
4: 6886 Chicago
1: 16716 Kansas
    [pem_pos=0, final_scores=0.5387687,
final_local_scores=0.51919496, pem=-0.3846871,
similarity=21.675514]
2: 70896 United_States_Department_of_Agriculture
    [pem_pos=0, final_scores=0.55574095,
final_local_scores=0.6366142, pem=0.0, similarity=23.087551]
3: 114560 Dodge_City,_Kansas
    [pem_pos=0, final_scores=0.27597922,
final_local_scores=0.27911633, pem=-0.13915609,
similarity=13.175854]
4: 6886 Chicago
    [pem_pos=0, final_scores=0.6672297,
final_local_scores=0.69994754, pem=-0.20658566,
similarity=26.100988]
```

Sample: b'582_Anti-abortion'

```
Anti-abortion speaker praises [1 1 Democrat 1] tolerance .
Alan Elsner [2 2 CHICAGO 2] 1996-08-27 An anti-abortion
politician addressed the [3 3 Democratic 3] convention on
Tuesday , but praised the overwhelmingly pro-abortion rights
party for its tolerance of his minority views . Rep . [4 4 To
ny Hall 4] of [5 5 Ohio 5] said he and other [6 6 Democrats 6]
```

who opposed abortion had always felt left out in their own party . " But this year is different . For the first time , the [7 7 Democratic Party 7] has included in our platform a conscience clause , " he said . The clause recognizes and welcomes [8 8 Democrats 8] with divergent views on abortion and states they have a full part to play at all levels of the party . " The [9 9 Democratic Party 9] is indeed the party of true inclusiveness , " [10 10 Hall 10] said . At its convention four years ago , organizers prevented then [11 11 Pennsylvania 11] Gov . [12 12 Robert Casey 12] , a vehement opponent of abortion , from speaking . [13 13 Republicans 13] have used their decision as an example of [14 14 Democrat 14] intolerance ever since . [15 15 Casey 15] told a news conference in [16 16 Chicago 16] on Tuesday he had asked to speak again this year but was turned down . [17 17 Democratic 17] leaders said there was not room on the program for every retired governor to speak . " I believe the [18 18 Democratic 18] party ought to be pro-woman , pro-child and pro-life , " [19 19 Casey 19] said . " I asked for the opportunity to deliver this message from the podium of the Democrat National Convention . For the second time in four years , my request fell on deaf ears , " he said . The [20 20 Republican Party 20] , whose platform calls for making all abortions illegal , faced a similar dilemma this year when [21 21 Massachusetts 21] Gov . [22 22 William Weld 22] asked to deliver a speech defending abortion rights and was turned down . [23 23 Bob Dole 23] , the [24 24 Republican 24] presidential nominee , tried and failed to insert a tolerance clause in his party 's platform recognizing the validity of those within the party who supported abortion rights . [25 25 Democrats 25] also heard Tuesday two passionate speeches defending abortion rights . [26 26 Kate Michelman 26] , president of the [27 27 National Abortion Rights Action League 27] , described how she had an abortion at a time when the procedure was illegal after her husband abandoned her with three young children . " I 'm here to speak up for choice and to speak for truth . The message from the [28 28 Republican Party 28] is one of disdain . Their answer to choice is control and punishment . Our answer is trust , compassion and respect , " she said . [29 29 Georgia 29] [30 30 Rep 30] . [31 31 Cynthia McKinney 31] said : " You make your moral decisions . I 'll make mine and let 's leave (Republican House Speaker) [32 32 Newt Gingrich 32] out of it . "

- 1: 5043544 Democratic_Party_(United_States)
- 2: 6886 Chicago
- 3: 5043544 Democratic_Party_(United_States)
- 4: 703518 Tony_P._Hall
- 5: 22199 Ohio
- 6: 5043544 Democratic_Party_(United_States)
- 7: 5043544 Democratic_Party_(United_States)
- 8: 5043544 Democratic_Party_(United_States)
- 9: 5043544 Democratic_Party_(United_States)

10: 703518 Tony_P._Hall
11: 23332 Pennsylvania
12: 619184 Robert_P._Casey
13: 32070 Republican_Party_(United_States)
14: 5043544 Democratic_Party_(United_States)
15: 619184 Robert_P._Casey
16: 6886 Chicago
17: 5043544 Democratic_Party_(United_States)
18: 5043544 Democratic_Party_(United_States)
19: 619184 Robert_P._Casey
20: 32070 Republican_Party_(United_States)
21: 1645518 Massachusetts
22: 505437 William_Weld
23: 55846 Bob_Dole
24: 32070 Republican_Party_(United_States)
25: 5043544 Democratic_Party_(United_States)
26: 14812059 Kate_Michelman
27: 770112 NARAL_Pro-Choice_America
28: 32070 Republican_Party_(United_States)
29: 48830 Georgia_(U.S._state)
30: 32070 Republican_Party_(United_States)
31: 392017 Cynthia_McKinney
32: 179187 Newt_Gingrich
1: 5043544 Democratic_Party_(United_States)
[pem_pos=0, final_scores=0.47253954,
final_local_scores=0.4056481, pem=-0.061652232,
similarity=16.530336]
2: 6886 Chicago
[pem_pos=0, final_scores=0.56623584,
final_local_scores=0.59980065, pem=-0.20658566,
similarity=23.112404]
3: 5043544 Democratic_Party_(United_States)
[pem_pos=0, final_scores=0.5098352,
final_local_scores=0.44638795, pem=-0.076093495,
similarity=17.824629]
4: 703518 Tony_P._Hall
[pem_pos=1, final_scores=0.514726,
final_local_scores=0.56324756, pem=-0.5098301,
similarity=23.67068]
5: 22199 Ohio
[pem_pos=0, final_scores=0.6772552,
final_local_scores=0.69839895, pem=-0.19147073,
similarity=25.972578]
6: 5043544 Democratic_Party_(United_States)
[pem_pos=0, final_scores=0.52580494,
final_local_scores=0.4638325, pem=-0.18007085,
similarity=18.910654]
7: 5043544 Democratic_Party_(United_States)
[pem_pos=0, final_scores=0.4552409,
final_local_scores=0.386752, pem=-0.33323824,
similarity=17.443365]
8: 5043544 Democratic_Party_(United_States)

```
[pem_pos=0, final_scores=0.53348404,
final_local_scores=0.47222072, pem=-0.18007085,
similarity=19.160975]
9: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.4770972,
final_local_scores=0.4106267, pem=-0.33323824,
similarity=18.155834]
10: 703518 Tony_P._Hall
    [pem_pos=1, final_scores=0.5701603,
final_local_scores=0.6238012, pem=-0.5098301,
similarity=25.47772]
11: 23332 Pennsylvania
    [pem_pos=0, final_scores=0.72125465,
final_local_scores=0.7490631, pem=-0.11810702,
similarity=27.085533]
12: 619184 Robert_P._Casey
    [pem_pos=0, final_scores=0.38815278,
final_local_scores=0.38206017, pem=0.0, similarity=15.49115]
13: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.46597463,
final_local_scores=0.39890373, pem=-0.11427376,
similarity=16.615234]
14: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.45541507,
final_local_scores=0.3869422, pem=-0.061652232,
similarity=15.9721155]
15: 619184 Robert_P._Casey
    [pem_pos=0, final_scores=0.8705471,
final_local_scores=0.91459984, pem=0.0, similarity=31.383196]
16: 6886 Chicago
    [pem_pos=0, final_scores=0.69123465,
final_local_scores=0.73634285, pem=-0.20658566,
similarity=27.187096]
17: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.49297726,
final_local_scores=0.4279732, pem=-0.076093495,
similarity=17.275097]
18: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.5643087,
final_local_scores=0.505892, pem=-0.076093495,
similarity=19.600348]
19: 619184 Robert_P._Casey
    [pem_pos=0, final_scores=0.98380667,
final_local_scores=1.0383186, pem=0.0, similarity=35.07521]
20: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.5155264,
final_local_scores=0.45303154, pem=-0.16706309,
similarity=18.517593]
21: 1645518 Massachusetts
    [pem_pos=0, final_scores=0.6594673,
final_local_scores=0.6748856, pem=-0.1423749,
similarity=25.0039]
```

```
22: 505437 William_Weld
    [pem_pos=0, final_scores=0.655134,
final_local_scores=0.66967756, pem=0.0, similarity=24.074226]
23: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.82006437,
final_local_scores=0.83845717, pem=0.0, similarity=29.110947]
24: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.49088436,
final_local_scores=0.42611384, pem=-0.1310191,
similarity=17.518303]
25: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.4619785,
final_local_scores=0.39411175, pem=-0.18007085,
similarity=16.830048]
26: 14812059 Kate_Michelman
    [pem_pos=0, final_scores=0.5082596,
final_local_scores=0.5907281, pem=0.0, similarity=21.718218]
27: 770112 NARAL_Pro-Choice_America
    [pem_pos=0, final_scores=0.38360363,
final_local_scores=0.40304118, pem=0.0, similarity=16.117266]
28: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.5152632,
final_local_scores=0.452744, pem=-0.16706309,
similarity=18.509012]
29: 48830 Georgia_(U.S._state)
    [pem_pos=0, final_scores=0.47261834,
final_local_scores=0.4965294, pem=-0.70047224,
similarity=22.716417]
30: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.3165924,
final_local_scores=0.2357263, pem=-0.9360621,
similarity=16.214708]
31: 392017 Cynthia_McKinney
    [pem_pos=0, final_scores=0.5422687,
final_local_scores=0.57633257, pem=0.0, similarity=21.288626]
32: 179187 Newt_Gingrich
    [pem_pos=0, final_scores=0.50652844,
final_local_scores=0.5133794, pem=0.0, similarity=19.409979]
```

Sample: b'583_U.S.'

```
[1 1 U.S. 1] Spring / White Wheat - Bids mostly steady . [2 2
CHICAGO 2] 1996-08-27 Dark northern spring and white wheat
bids were mostly steady on Tuesday but a few locations quoted
weaker values as newly harvested spring wheat flooded the
market , several cash grain dealers said . " There 's too
much nearby wheat coming into the market so we 're backing
off the basis to slow it down , " a [3 3 Montana 3] dealer
said . Bids there dropped 10 cents per bushel . Harvest was
also progressing well in parts of [4 4 North Dakota 4] , but
one dealer there said new crop movement remained limited to a
steady trickle . " We 're seeing some new crop coming in now
but it 's slow going , " the dealer said . Elsewhere , basis
```

values were mostly steady in quiet conditions with little noteworthy domestic or export business , dealers said . Durum bids were steady after jumping 50 cents per bushel in some areas on Monday . Price per bushel for 14-pct protein dark northern spring , durum and white wheats , in dollars per bushel : Spring Chg Durum (m) Chg White Chg MINNESOTA Minneapolis 5.06 up .02 5.75 unc -- -- [5 5 Duluth 5] 5.06 up .02 --- --- NORTH DAKOTA Hunter ([6 6 Red River 6]) 4.46 dn .02 5.00 unc -- -- Spring Chg HRW 12pct Chg White Chg [7 7 Billings 7] MT 4.62 up .01 4.50 dn .01 --- -- [8 8 Havre 8] MT 4.54 dn .10 --- --- [9 9 Rudyard 9] MT 4.54 dn .10 --- --- [10 10 Wolf Point 10] MT 4.41 dn .10 --- --- [11 11 Portland 11] OR 5.60 up .02 5.1700 dn .01 [12 12 Pendleton 12] OR --- --- 4.7300 up .01 Coolee City WA 5.13 up .02 --- 4.7000 unc [13 13 Waterville 13] WA 5.05 up .02 --- 4.6200 unc [14 14 Wenatchee 14] WA 5.15 up .02 --- 4.7200 unc note : nc=acomparison , na=not available ([15 15 Chicago 15] bureau 312-408-8720)

1: 3434750 United_States

2: 6886 Chicago

3: 19978 Montana

4: 21651 North_Dakota

5: 63503 Duluth,_Minnesota

6: 150045 Red_River_of_the_North

7: 123654 Billings,_Montana

8: 123412 Havre,_Montana

9: 123438 Rudyard,_Montana

10: 123602 Wolf_Point,_Montana

11: 23503 Portland,_Oregon

12: 130965 Pendleton,_Oregon

13: 137900 Waterville,_Washington

14: 137842 Wenatchee,_Washington

15: 6886 Chicago

1: 3434750 United_States

[pem_pos=0, final_scores=0.484991,

final_local_scores=0.48399252, pem=-0.14075686,

similarity=19.298473]

2: 6886 Chicago

[pem_pos=0, final_scores=0.6602648,

final_local_scores=0.6577056, pem=-0.20658566,

similarity=24.840406]

3: 19978 Montana

[pem_pos=0, final_scores=0.55115145,

final_local_scores=0.53632545, pem=-0.308461,

similarity=21.772194]

4: 21651 North_Dakota

[pem_pos=0, final_scores=0.47323436,

final_local_scores=0.43865597, pem=-0.33139732,

similarity=18.982271]

5: 63503 Duluth,_Minnesota

[pem_pos=0, final_scores=0.35356385,

```
final_local_scores=0.31763202, pem=-0.2524106,
similarity=14.9411335]
6: 150045 Red_River_of_the_North
    [pem_pos=1, final_scores=0.2136506,
final_local_scores=0.2376681, pem=-1.4132268,
similarity=18.86755]
7: 123654 Billings,_Montana
    [pem_pos=0, final_scores=0.4064321,
final_local_scores=0.3634767, pem=-0.2673195,
similarity=16.390306]
8: 123412 Havre,_Montana
    [pem_pos=0, final_scores=0.40305525,
final_local_scores=0.37876344, pem=0.0, similarity=15.39277]
9: 123438 Rudyard,_Montana
    [pem_pos=0, final_scores=0.37020904,
final_local_scores=0.47034675, pem=-1.0018251,
similarity=23.573877]
10: 123602 Wolf_Point,_Montana
    [pem_pos=0, final_scores=0.46695703,
final_local_scores=0.48045462, pem=0.0, similarity=18.427437]
11: 23503 Portland,_Oregon
    [pem_pos=0, final_scores=0.455653,
final_local_scores=0.42396134, pem=-0.66320467,
similarity=20.348175]
12: 130965 Pendleton,_Oregon
    [pem_pos=0, final_scores=0.5188969,
final_local_scores=0.5081377, pem=-1.4213227,
similarity=26.98293]
13: 137900 Waterville,_Washington
    [pem_pos=3, final_scores=0.4363914,
final_local_scores=0.474724, pem=-1.9642596,
similarity=28.938372]
14: 137842 Wenatchee,_Washington
    [pem_pos=0, final_scores=0.39879078,
final_local_scores=0.39036512, pem=-0.0781071,
similarity=16.163746]
15: 6886 Chicago
    [pem_pos=0, final_scores=0.6852587,
final_local_scores=0.68500763, pem=-0.20658566,
similarity=25.655153]
```

Sample: b'584_Birmingham'
Birmingham Public Park , [1 1 Ala 1] . , Aa3 / VMIG-1 -
Moody 's . [2 2 NEW YORK 2] 1996-08-27 Moody 's Investors
Service - Rating Announcement As of 08/23/96 . Issuer :
Birmingham Public Park & Rec . Bd . Revenue ref . ([3 3 YMCA
3] Proj .) ser . ' 96 State : AL Rating : Aa3 / VMIG 1 Sale
Amount : 3,390,000 Expected Sale Date : 08/28/96 -- U.S.
Municipal Desk , 212-859-1650

1: 303 Alabama
2: 645042 New_York_City
3: 87241 YMCA

```
1: 303 Alabama
    [pem_pos=7, final_scores=0.6427662,
final_local_scores=0.6798346, pem=-2.937321,
similarity=40.35094]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.45451695,
final_local_scores=0.4447037, pem=-1.7906687,
similarity=27.098494]
3: 87241 YMCA
    [pem_pos=0, final_scores=0.49579173,
final_local_scores=0.5060461, pem=-0.022545643,
similarity=19.313745]
```

Sample: b'585_U.S.'

[1 1 U.S. 1] lauds Russian-Chechen deal . [2 2 WASHINGTON 2]
1996-08-27 The [3 3 United States 3] on Tuesday welcomed a
deal aimed at resuming a troop withdrawal from the embattled [
4 4 Chechen 4] capital , [5 5 Grozny 5] . " That is a welcome
development . We urge both sides to continue their dialogue
aimed at reaching a political settlement " of the 20-month
conflict between [6 6 Russian 6] troops and [7 7 Chechen 7]
rebels , [8 8 State Department 8] spokesman Glyn Davies
said . The commander of [9 9 Russian 9] troops in [10 10 Chec
hnyia 10] , Vyacheslav Tikhomirov , and [11 11 Chechen 11]
rebel chief-of-staff [12 12 Aslan Mashadov 12] signed the
deal under which the troop withdrawal is to resume on
Wednesday .

```
1: 3434750 United_States
2: 108956 Washington,_D.C.
3: 3434750 United_States
4: 6095 Chechnya
5: 140790 Grozny
6: 25391 Russia
7: 6095 Chechnya
8: 31975 United_States_Department_of_State
9: 25391 Russia
10: 6095 Chechnya
11: 6095 Chechnya
12: 340963 Aslan_Mashadov
1: 3434750 United_States
    [pem_pos=0, final_scores=0.44759172,
final_local_scores=0.45274585, pem=-0.14075686,
similarity=18.366009]
2: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.5540675,
final_local_scores=0.5899283, pem=-1.7496481,
similarity=31.209208]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.4194523,
final_local_scores=0.4220078, pem=-0.073457584,
similarity=17.08274]
4: 6095 Chechnya
```

```
[pem_pos=2, final_scores=0.4999507,
final_local_scores=0.4557621, pem=-0.8617233,
similarity=22.376749]
5: 140790 Grozny
[pem_pos=0, final_scores=0.51126164,
final_local_scores=0.5207832, pem=0.0, similarity=19.630922]
6: 25391 Russia
[pem_pos=1, final_scores=0.8944705,
final_local_scores=0.92290443, pem=-1.0515022,
similarity=37.349255]
7: 6095 Chechnya
[pem_pos=2, final_scores=0.6099232,
final_local_scores=0.5758903, pem=-0.8617233,
similarity=25.961615]
8: 31975 United_States_Department_of_State
[pem_pos=0, final_scores=0.42859477,
final_local_scores=0.4440379, pem=0.0, similarity=17.34069]
9: 25391 Russia
[pem_pos=1, final_scores=0.8638688,
final_local_scores=0.8894766, pem=-1.0515022,
similarity=36.351704]
10: 6095 Chechnya
[pem_pos=0, final_scores=0.58470064,
final_local_scores=0.5483385, pem=0.0, similarity=20.453226]
11: 6095 Chechnya
[pem_pos=2, final_scores=0.54528236,
final_local_scores=0.50528, pem=-0.8617233,
similarity=23.854462]
12: 340963 Aslan_Maskhadov
[pem_pos=0, final_scores=0.60475546,
final_local_scores=0.65188986, pem=0.0, similarity=23.543407]
```

Sample: b'586_Akron'

```
[1 1 Akron 1] , [2 2 Ohio 2] , $ 6 mln bonds rated single-A
- Moody 's . [3 3 NEW YORK 3] , Aug 27 - Moody 's Investors
Service - Rating Announcement As of 08/26/96 . Issuer : [4 4 A
kron 4] State : OH Rating : A Sale Amount : 6,310,000
Expected Sale Date : 08/28/96
1: 129842 Akron,_Ohio
2: 22199 Ohio
3: 645042 New_York_City
4: 129842 Akron,_Ohio
1: 129842 Akron,_Ohio
[pem_pos=0, final_scores=0.47999138,
final_local_scores=0.42053503, pem=-0.6020413,
similarity=19.913311]
2: 22199 Ohio
[pem_pos=0, final_scores=0.655258,
final_local_scores=0.6371724, pem=-0.19147073,
similarity=24.145456]
3: 645042 New_York_City
[pem_pos=1, final_scores=0.4126504,
```

```
final_local_scores=0.39989442, pem=-1.7906687,
similarity=25.761295]
4: 129842 Akron,_Ohio
    [pem_pos=0, final_scores=0.422769,
final_local_scores=0.35802832, pem=-0.6020413,
similarity=18.047987]

Sample: b'587_Stallone'
[1 1 Stallone 1] , fiancee have baby girl . [2 2 MIAMI 2]
1996-08-27 Actor [3 3 Sylvester Stallone 3] and his fiancee ,
model [4 4 Jennifer Flavin 4] , had a baby girl on Tuesday , [
5 5 Stallone 5] 's publicist said . The 7-pound , 4-ounce
( 3.3 kg ) girl , named Sophia Rose , was born shortly after
midnight at South Miami Hospital , publicist Paul Bloch
said . " Both mother and baby are doing fine and are in
wonderful health , " he said , adding that it was the couple
's first child . He said [6 6 Stallone 6] , best known for
the " [7 7 Rocky 7] " and " [8 8 Rambo 8] " movies , left the
set of " Copland , " which is filming in [9 9 New York 9] and
[10 10 New Jersey 10] , to be with [11 11 Flavin 11] for the
birth .
1: 61404 Sylvester_Stallone
2: 53846 Miami
3: 61404 Sylvester_Stallone
4: 1045992 Jennifer_Flavin
5: 61404 Sylvester_Stallone
6: 61404 Sylvester_Stallone
7: 45772 Rocky
8: 380769 Rambo_(film_series)
9: 645042 New_York_City
10: 21648 New_Jersey
11: 1045992 Jennifer_Flavin
1: 61404 Sylvester_Stallone
    [pem_pos=0, final_scores=0.36377507,
final_local_scores=0.318809, pem=0.0, similarity=13.603609]
2: 53846 Miami
    [pem_pos=0, final_scores=0.49469274,
final_local_scores=0.51943475, pem=-0.35396606,
similarity=21.515604]
3: 61404 Sylvester_Stallone
    [pem_pos=0, final_scores=0.5126658,
final_local_scores=0.48144954, pem=0.0, similarity=18.457129]
4: 1045992 Jennifer_Flavin
    [pem_pos=0, final_scores=0.512806,
final_local_scores=0.568703, pem=0.0, similarity=21.060946]
5: 61404 Sylvester_Stallone
    [pem_pos=0, final_scores=0.37638634,
final_local_scores=0.33258492, pem=0.0, similarity=14.01471]
6: 61404 Sylvester_Stallone
    [pem_pos=0, final_scores=0.45515817,
final_local_scores=0.4186312, pem=0.0, similarity=16.582502]
7: 45772 Rocky
```

```
[pem_pos=0, final_scores=0.40879673,
final_local_scores=0.42029768, pem=-0.4167528,
similarity=18.898602]
8: 380769 Rambo_(film_series)
    [pem_pos=3, final_scores=0.13364999,
final_local_scores=0.159774, pem=-2.1562705,
similarity=20.583822]
9: 645042 New_York_City
    [pem_pos=1, final_scores=0.41427022,
final_local_scores=0.43566877, pem=-1.7906687,
similarity=26.828873]
10: 21648 New_Jersey
    [pem_pos=0, final_scores=0.34649134,
final_local_scores=0.40163422, pem=-0.13504621,
similarity=16.80968]
11: 1045992 Jennifer_Flavin
    [pem_pos=0, final_scores=0.83731705,
final_local_scores=0.9520586, pem=0.0, similarity=32.50104]
```

Sample: b'588_Poll'

Poll shows [1 1 Clinton 1] lead over [2 2 Dole 2] jumps to 15 pts . [3 3 CHICAGO 3] 1996-08-27 An [4 4 ABC News 4] poll released on Tuesday showed President [5 5 Bill Clinton 5] 's lead over [6 6 Republican 6] challenger [7 7 Bob Dole 7] stretching to 15 points in advance of the Nov. 5 election . The poll , taken on Sunday and Monday as the president engaged in a whistle-stop train trip to the [8 8 Democratic Co nvention 8] in [9 9 Chicago 9] , put [10 10 Clinton 10] at 51 percent , [11 11 Dole 11] at 36 percent and [12 12 Ross Perot 12] of the [13 13 Reform Party 13] at 8 percent . A similar poll conducted on Saturday and Sunday had showed a nine point lead for [14 14 Clinton 14] , ahead by 47-38 percent . [15 15 Dole 15] , down by around 20 points in early August in [16 16 ABC 16] polls , had closed to within four percentage points immediately after the [17 17 Republican 17] convention in [18 18 San Diego 18] earlier in August . Other polls also showed a strong [19 19 Dole 19] bounce after [20 20 San Diego 20] but [21 21 Clinton 21] then rebuilding his lead . Tuesday 's poll involved 1,002 registered voters and had a margin of error of 3.5 percentage points . [22 22 ABC 22] said a parallel poll of 824 likely voters showed the president ahead by 11 points , with [23 23 Clinton 23] at 50 percent , [24 24 Dole 24] at 39 percent and [25 25 Perot 25] at 6 percent . The poll of registered voters showed a shift in favor of the [26 26 Democrats 26] in the elections for [27 27 House of Repr esentatives 27] , with 51 percent saying that if the vote were today they would go for a [28 28 Democrat 28] and 41 percent opting for a [29 29 Republican 29] . That compared with the previous poll 's 48-43 lead for the [30 30 Democrats 30] . The poll gave [31 31 Clinton 31] a 53 percent to 39 percent lead over [32 32 Dole 32] if [33 33 Perot 33] were not in the race . The poll indicated a fall in the number of

people who believed [34 34 Dole 34] would be able to fulfil his promise to cut the federal budget deficit and cut income taxes by 15 percent at the same time . It showed 23 percent believed it possible compared to 70 who believed it was n't . That compared to 26-57 in the previous poll .

1: 3356 Bill_Clinton
2: 55846 Bob_Dole
3: 6886 Chicago
4: 62027 American_Broadcasting_Company
5: 3356 Bill_Clinton
6: 32070 Republican_Party_(United_States)
7: 55846 Bob_Dole
8: 468102 Democratic_National_Convention
9: 6886 Chicago
10: 3356 Bill_Clinton
11: 55846 Bob_Dole
12: 43505 Ross_Perot
13: 295504 Reform_Party_of_the_United_States_of_America
14: 3356 Bill_Clinton
15: 55846 Bob_Dole
16: 62027 American_Broadcasting_Company
17: 32070 Republican_Party_(United_States)
18: 28504 San_Diego
19: 55846 Bob_Dole
20: 28504 San_Diego
21: 3356 Bill_Clinton
22: 62027 American_Broadcasting_Company
23: 3356 Bill_Clinton
24: 55846 Bob_Dole
25: 43505 Ross_Perot
26: 5043544 Democratic_Party_(United_States)
27: 19468532 House_of_Representatives
28: 5043544 Democratic_Party_(United_States)
29: 32070 Republican_Party_(United_States)
30: 5043544 Democratic_Party_(United_States)
31: 3356 Bill_Clinton
32: 55846 Bob_Dole
33: 43505 Ross_Perot
34: 55846 Bob_Dole
1: 3356 Bill_Clinton
 [pem_pos=0, final_scores=0.8246383,
final_local_scores=0.8084028, pem=0.0, similarity=28.214064]
2: 55846 Bob_Dole
 [pem_pos=0, final_scores=0.8398605,
final_local_scores=0.8185185, pem=0.0, similarity=28.515938]
3: 6886 Chicago
 [pem_pos=0, final_scores=0.62465954,
final_local_scores=0.67137176, pem=-0.20658566,
similarity=25.24823]
4: 62027 American_Broadcasting_Company
 [pem_pos=2, final_scores=0.18323283,
final_local_scores=0.20833763, pem=-0.68394935,

```
similarity=14.026346]
5: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.59806985,
final_local_scores=0.5592837, pem=0.0, similarity=20.779854]
6: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.43695378,
final_local_scores=0.40079248, pem=-0.1310191,
similarity=16.762663]
7: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.7503482,
final_local_scores=0.71939784, pem=0.0, similarity=25.55798]
8: 468102 Democratic_National_Convention
    [pem_pos=0, final_scores=0.40784442,
final_local_scores=0.3854018, pem=-0.23190276,
similarity=16.851994]
9: 6886 Chicago
    [pem_pos=0, final_scores=0.64375424,
final_local_scores=0.6922298, pem=-0.20658566,
similarity=25.870676]
10: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.8338664,
final_local_scores=0.8184832, pem=0.0, similarity=28.514883]
11: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.8522287,
final_local_scores=0.8320288, pem=0.0, similarity=28.919111]
12: 43505 Ross_Perot
    [pem_pos=0, final_scores=0.5082758,
final_local_scores=0.48732424, pem=0.0, similarity=18.63244]
13: 295504 Reform_Party_of_the_United_States_of_America
    [pem_pos=0, final_scores=0.48685467,
final_local_scores=0.49718517, pem=-0.8355943,
similarity=23.4708]
14: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.81906086,
final_local_scores=0.8023103, pem=0.0, similarity=28.032251]
15: 55846 Bob_Dole
    [pem_pos=0, final_scores=1.005838,
final_local_scores=0.99982375, pem=0.0, similarity=33.92645]
16: 62027 American_Broadcasting_Company
    [pem_pos=0, final_scores=0.5078664,
final_local_scores=0.56295043, pem=-0.3336378,
similarity=22.703651]
17: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.46575588,
final_local_scores=0.43225437, pem=-0.1310191,
similarity=17.701548]
18: 28504 San_Diego
    [pem_pos=0, final_scores=0.4462903,
final_local_scores=0.5041517, pem=-0.18777876,
similarity=20.155775]
19: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.7751904,
```

```
final_local_scores=0.7478762, pem=0.0, similarity=26.407831]
20: 28504 San_Diego
    [pem_pos=0, final_scores=0.5013773,
final_local_scores=0.56432587, pem=-0.18777876,
similarity=21.951492]
21: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.8293362,
final_local_scores=0.8135346, pem=0.0, similarity=28.367208]
22: 62027 American_Broadcasting_Company
    [pem_pos=0, final_scores=0.40482658,
final_local_scores=0.4503951, pem=-0.3336378,
similarity=19.344776]
23: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.81432444,
final_local_scores=0.7971364, pem=0.0, similarity=27.877853]
24: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.830857,
final_local_scores=0.8086835, pem=0.0, similarity=28.222443]
25: 43505 Ross_Perot
    [pem_pos=0, final_scores=0.45585048,
final_local_scores=0.43282372, pem=0.0, similarity=17.006037]
26: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.48157454,
final_local_scores=0.4514591, pem=-0.18007085,
similarity=18.541407]
27: 19468532 House_of_Representatives
    [pem_pos=5, final_scores=0.23076214,
final_local_scores=0.32193917, pem=-3.0198689,
similarity=30.119535]
28: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.4240275,
final_local_scores=0.38859773, pem=-0.061652232,
similarity=16.021519]
29: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.419424,
final_local_scores=0.3816439, pem=-0.1310191,
similarity=16.19123]
30: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.48956102,
final_local_scores=0.46018308, pem=-0.18007085,
similarity=18.801748]
31: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.7949286,
final_local_scores=0.77594954, pem=0.0, similarity=27.245594]
32: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.9449573,
final_local_scores=0.9333208, pem=0.0, similarity=31.941868]
33: 43505 Ross_Perot
    [pem_pos=0, final_scores=0.5165753,
final_local_scores=0.4991563, pem=0.0, similarity=18.985533]
34: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.9049993,
```

```
final_local_scores=0.8896727, pem=0.0, similarity=30.639322]
```

Sample: b'589_U.S.'

[1 1 U.S. 1] bulk millfeeds - Immediate supply tight . [2 2 C
HICAGO 2] 1996-08-27 Millfeed supplies for prompt shipment remained tight and prices continued to move higher , millfeed dealers said . High-priced corn and increased demand for livestock feed continued to support millfeed prices in nearly all sectors . Flour mills sold much of their production through September leaving little available for prompt shipment . [3 3 Portland 3] sources said feed mixer demand was keeping pace with millfeed production and driving prices higher . [4 4 Portland 4] sources said with corn priced there at \$ 200 per ton and barley at \$ 140 , the millfeeds at \$ 125 represent a good value . In the southeast [5 5 U.S. 5] , dealers said feed mixers continued to be steady buyers with demand increasing for October to March positions . The closely-watched [6 6 Kansas City 6] rail market was steady at \$ 115 per ton bid and \$ 118 offered . -- [7 7 Chicago 7]

newsdesk 312-408-8720--

1: 3434750 United_States
2: 6886 Chicago
3: 23503 Portland,_Oregon
4: 23503 Portland,_Oregon
5: 3434750 United_States
6: 17454 Kansas_City,_Missouri
7: 6886 Chicago

1: 3434750 United_States
[pem_pos=0, final_scores=0.5152087,
final_local_scores=0.4865151, pem=-0.14075686,
similarity=19.373753]

2: 6886 Chicago
[pem_pos=0, final_scores=0.717535,
final_local_scores=0.6747952, pem=-0.20658566,
similarity=25.350393]

3: 23503 Portland,_Oregon
[pem_pos=0, final_scores=0.47849476,
final_local_scores=0.42437845, pem=-0.66320467,
similarity=20.360622]

4: 23503 Portland,_Oregon
[pem_pos=0, final_scores=0.4792934,
final_local_scores=0.4252509, pem=-0.66320467,
similarity=20.386658]

5: 3434750 United_States
[pem_pos=0, final_scores=0.49505383,
final_local_scores=0.46449894, pem=-0.14075686,
similarity=18.716745]

6: 17454 Kansas_City,_Missouri
[pem_pos=0, final_scores=0.2875306,
final_local_scores=0.22503299, pem=-0.8046767,
similarity=15.181103]

7: 6886 Chicago

```
[pem_pos=0, final_scores=0.72676796,
final_local_scores=0.6848808, pem=-0.20658566,
similarity=25.651367]
```

Sample: b'590_Puerto'

[1 1 Puerto Rico 1] girl has surgery for hairy face . [2 2 PH
ILADELPHIA 2] 1996-08-27 A two-year old [3 3 Puerto Rican 3]
girl began surgical treatment on Tuesday for a rare condition
that has left half of her face covered with a hairy , dark-
brown patch of skin . The girl , Abyss DeJesus , suffers from
a " hairy nevus " on the right side of her face , a condition
that has only been reported a few times in medical journals ,
the St. Christopher Children 's Hospital said . In addition
to social ostracism , the condition also carries a high risk
of cancer . It will be corrected by gradually expanding
healthy skin with a surgical balloon , then transplanting
that skin to the afflicted side of her face . " She is doing
well , " hospital spokeswoman Carol Norris said . " The
surgery is under way . " Norris said Tuesday 's surgery
involved placing five balloons in DeJesus 's forehead ,
shoulders and the back of her neck and partially filling them
with a saline solution . More saline solution will be
inserted in 16 weekly treatments . The girl , who was
accompanied to [4 4 Philadelphia 4] by her parents , will
need more surgery later to correct the condition on her
chest , back and legs , the hospital said .

1: 23041 Puerto_Rico

2: 50585 Philadelphia

3: 23041 Puerto_Rico

4: 50585 Philadelphia

1: 23041 Puerto_Rico

```
[pem_pos=0, final_scores=0.48015046,
final_local_scores=0.45764124, pem=-0.2967188,
similarity=19.360243]
```

2: 50585 Philadelphia

```
[pem_pos=0, final_scores=0.59191364,
final_local_scores=0.5797255, pem=-0.2107264,
similarity=22.535841]
```

3: 23041 Puerto_Rico

```
[pem_pos=0, final_scores=0.3707283,
final_local_scores=0.33811414, pem=-0.5221348,
similarity=17.019163]
```

4: 50585 Philadelphia

```
[pem_pos=0, final_scores=0.5503388,
final_local_scores=0.53431135, pem=-0.2107264,
similarity=21.180592]
```

Sample: b'591_Eau'

[1 1 Eau Claire 1] , 1 Wisc . revs won by [2 2 Robert W. Baird 2] . [3 3 NEW YORK 3] 1996-08-27 Robert W. Baird & Co . , Inc. , said it won \$ 1 million of [4 4 Eau Claire 4] , 2 Wisc . , waterworks system mortgage revenue bonds , Series

1996 , with a true interest cost of 5.2893 percent .

1: 151296 Eau_Claire,_Wisconsin
2: 12327134 Robert_W._Baird
3: 645042 New_York_City
4: 151296 Eau_Claire,_Wisconsin
1: 33127 Wisconsin
[winner: 2165061 WISC-TV, pem_pos=0,
final_scores=-0.20998709, final_local_scores=-0.15663208,
pem=0.0, similarity=-0.58450073]
RECALL MISS: 33127 Wisconsin]
[winner: 330842 Achang_people, pem_pos=1,
final_scores=-1.3940604, final_local_scores=-1.377712,
pem=-2.4264104, similarity=-23.828768]
RECALL MISS: 33127 Wisconsin]
2: 33127 Wisconsin
[winner: 2165061 WISC-TV, pem_pos=0,
final_scores=-0.3005268, final_local_scores=-0.25553298,
pem=0.0, similarity=-3.5359004]
RECALL MISS: 33127 Wisconsin]
[winner: 4721653 Zyphe_language, pem_pos=12,
final_scores=-1.4990907, final_local_scores=-1.4347568,
pem=-4.4188404, similarity=-14.69595]
RECALL MISS: 33127 Wisconsin]
1: 151296 Eau_Claire,_Wisconsin
[pem_pos=0, final_scores=0.3868683,
final_local_scores=0.3526012, pem=-0.57503253,
similarity=17.739153]
2: 12327134 Robert_W._Baird
[pem_pos=0, final_scores=0.35313642,
final_local_scores=0.47003275, pem=0.0, similarity=18.116428]
3: 645042 New_York_City
[pem_pos=1, final_scores=0.421654,
final_local_scores=0.44811267, pem=-1.7906687,
similarity=27.200224]
4: 151296 Eau_Claire,_Wisconsin
[pem_pos=0, final_scores=0.3796935,
final_local_scores=0.34476385, pem=-0.57503253,
similarity=17.50527]

Sample: b'592_Massachusetts'

[1 1 Massachusetts 1] home sales dip in July - report . [2 2 BOSTON 2] 1996-08-27 Home sales across [3 3 Massachusetts 3] were down 2.3 percent in July , compared to a month earlier , but up 21 percent for the year , according to the Massachusetts Association of Realtors . The association said a total of 4,464 single-family homes were sold in July , compared to 4,570 in June . The average selling price , \$ 206,464 , was up 10.6 percent over July 1995 . Condominium sales edged up 6.0 percent for July and 24.8 percent for the year , the group said , while prices for condos nudged up less than 1.0 percent to an average of \$ 123,394 . In July , the average rate on a 30-year fixed rate mortgage was 8.25

percent , below June 's 8.32 percent but still higher than February 's 7.03 percent , the report noted . -- [4 4 Boston 4] bureau , 617-367-4106

1: 1645518 Massachusetts

2: 24437894 Boston

3: 1645518 Massachusetts

4: 24437894 Boston

1: 1645518 Massachusetts

[pem_pos=0, final_scores=0.6755483, final_local_scores=0.62750363, pem=-0.1423749, similarity=23.58993]

2: 24437894 Boston

[pem_pos=0, final_scores=0.7172285, final_local_scores=0.67303276, pem=-0.23164093, similarity=25.434052]

3: 1645518 Massachusetts

[pem_pos=0, final_scores=0.6415061, final_local_scores=0.5903176, pem=-0.1423749, similarity=22.480225]

4: 24437894 Boston

[pem_pos=0, final_scores=0.65558356, final_local_scores=0.6056951, pem=-0.23164093, similarity=23.424562]

Sample: b'593_Amtrak'

[1 1 Amtrak 1] train derails , three injured - officials . [2 2 MONTPELIER 2] , Vt . 1996-08-27 At least three people were injured when an [3 3 Amtrak 3] passenger train slammed into an empty logging truck and derailed Tuesday , officials said . The Vermonter , which runs between [4 4 St. Albans 4] , Vermont , near the [5 5 Canadian 5] border and 1 Washington , D.C. , collided with the truck at 7:51 a.m. EDT near the rural town of [6 6 Roxbury 6] some 15 miles southeast of the state capital [7 7 Montpelier 7] , [8 8 Amtrak 8] spokeswoman Maureen Garrity said . Vermont Central Hospital spokesman Dan Pudvah said two of the injured were treated there -- the truck driver , who was suffering from multiple trauma injuries , and a passenger . Pudvah said he understood other people with minor injuries were being treated at the scene . Garrity said a train conductor was also injured . The train 's engine and its six cars derailed but were still standing , state police said . The exact number of passangers on the train was not known . " We had 70 reservations for the train , but that does n't mean there were 70 passengers aboard , " Garrity said . Uninjured passengers were to be taken by bus to [9 9 Springfield 9] , [10 10 Massachusetts 10] , where they will be put aboard another train to continue their journey to [11 11 New York City 11] and [12 12 Washington 12] , Garrity said . She said the train was travelling at 54 mph when it crashed into the truck , which was crossing the tracks onto a dirt road in the rural area bordering the Northfield Mountains .

```
1: 51928 Amtrak
2: 41992 Montpelier,_Vermont
3: 51928 Amtrak
4: 260221 St._Albans_(city),_Vermont
5: 5042916 Canada
6: 137418 Roxbury,_Vermont
7: 41992 Montpelier,_Vermont
8: 51928 Amtrak
9: 92402 Springfield,_Massachusetts
10: 1645518 Massachusetts
11: 645042 New_York_City
12: 108956 Washington,_D.C.
1: 108956 Washington,_D.C.
    [winner: 108956 Washington,_D.C., pem_pos=0,
final_scores=-0.11020459, final_local_scores=-0.19383106,
pem=-1.7496481, similarity=7.8202653]
    gt: 108956 Washington,_D.C., pem_pos=0,
final_scores=-0.11020459, final_local_scores=-0.19383106,
pem=-1.7496481, similarity=7.8202653 ]
    [winner: 108956 Washington,_D.C., pem_pos=0,
final_scores=-0.18982802, final_local_scores=-0.28080755,
pem=-0.15729709, similarity=-3.4347384]
    gt: 108956 Washington,_D.C., pem_pos=0,
final_scores=-0.18982802, final_local_scores=-0.28080755,
pem=-0.15729709, similarity=-3.4347384 ]
    [winner: 37537378 Admissions_(2011_film), pem_pos=12,
final_scores=-1.8786082, final_local_scores=-1.8695264,
pem=-4.6728287, similarity=-26.289114]
    RECALL MISS: 108956 Washington,_D.C.]
1: 51928 Amtrak
    [pem_pos=0, final_scores=0.36882836,
final_local_scores=0.36287656, pem=0.0, similarity=14.9186735]
2: 41992 Montpelier,_Vermont
    [pem_pos=0, final_scores=0.42890698,
final_local_scores=0.39990056, pem=-0.89688367,
similarity=20.900936]
3: 51928 Amtrak
    [pem_pos=0, final_scores=0.3224753,
final_local_scores=0.3122429, pem=0.0, similarity=13.407663]
4: 260221 St._Albans_(city),_Vermont
    [pem_pos=0, final_scores=0.35929734,
final_local_scores=0.332351, pem=-1.4735126,
similarity=22.02092]
5: 5042916 Canada
    [pem_pos=0, final_scores=0.5851148,
final_local_scores=0.6203361, pem=-0.41303828,
similarity=24.847946]
6: 137418 Roxbury,_Vermont
    [pem_pos=4, final_scores=0.24830215,
final_local_scores=0.3033959, pem=-2.104593,
similarity=24.588757]
7: 41992 Montpelier,_Vermont
```

```
[pem_pos=0, final_scores=0.408925,
final_local_scores=0.37807333, pem=-0.89688367,
similarity=20.249569]
8: 51928 Amtrak
    [pem_pos=0, final_scores=0.34727734,
final_local_scores=0.33933538, pem=0.0, similarity=14.216158]
9: 92402 Springfield,_Massachusetts
    [pem_pos=0, final_scores=0.22413908,
final_local_scores=0.16724709, pem=-1.8459358,
similarity=19.119186]
10: 1645518 Massachusetts
    [pem_pos=0, final_scores=0.51764864,
final_local_scores=0.49002713, pem=-0.1423749,
similarity=19.487356]
11: 645042 New_York_City
    [pem_pos=0, final_scores=0.611639,
final_local_scores=0.60530657, pem=0.0, similarity=22.153269]
12: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.50039667,
final_local_scores=0.48478442, pem=-1.7496481,
similarity=28.071505]
```

Sample: b'594_Paralympics'

[1 1 Paralympics 1] an example for gloomy France-Juppe . [2 2 PARIS 2] 1996-08-27 Prime Minister [3 3 Alain Juppe 3] on Tuesday hailed handicapped athletes who took part in [4 4 Atlanta 4] 's [5 5 Paralympic Games 5] as an example for gloom-stricken [6 6 France 6] . " What we hear every morning is gloom , resignation and scepticism ... You are the opposite , " Juppe told a successful [7 7 French 7] team at [8 8 Paris 8] airport as he welcomed them back from the games which followed the July-August [9 9 Olympics 9] . " If you had been struck ... by the disease of scepticism , gloom and resignation , you would not be here . You are a true example for the nation , " he said . The [10 10 French 10] team won 95 medals in [11 11 Atlanta 11] , 35 of them gold . Opinion polls consistently show [12 12 French 12] voters pessimistic and fed up as the economy stagnates and unemployment lingers at near-record levels .

```
1: 67537 Paralympic_Games
2: 22989 Paris
3: 434931 Alain_Juppé
4: 3138 Atlanta
5: 67537 Paralympic_Games
6: 5843419 France
7: 5843419 France
8: 22989 Paris
9: 72311 1996_Summer_Olympics
10: 5843419 France
11: 3138 Atlanta
12: 5843419 France
1: 67537 Paralympic_Games
```

```
[pem_pos=0, final_scores=0.3081059,
final_local_scores=0.33883244, pem=-0.36854455,
similarity=16.20535]
2: 22989 Paris
    [pem_pos=0, final_scores=0.61275727,
final_local_scores=0.5931495, pem=-0.115301244,
similarity=22.417503]
3: 434931 Alain_Juppé
    [pem_pos=0, final_scores=0.6066051,
final_local_scores=0.7406213, pem=0.0, similarity=26.19133]
4: 3138 Atlanta
    [pem_pos=0, final_scores=0.42962855,
final_local_scores=0.42815894, pem=-0.23152597,
similarity=18.125904]
5: 67537 Paralympic_Games
    [pem_pos=0, final_scores=0.2837687,
final_local_scores=0.3122478, pem=-0.10437569,
similarity=13.975422]
6: 5843419 France
    [pem_pos=0, final_scores=0.67601615,
final_local_scores=0.6475626, pem=-0.16240834,
similarity=24.297474]
7: 5843419 France
    [pem_pos=0, final_scores=0.8016494,
final_local_scores=0.78479785, pem=-0.80223274,
similarity=31.872313]
8: 22989 Paris
    [pem_pos=0, final_scores=0.67352086,
final_local_scores=0.6595245, pem=-0.115301244,
similarity=24.398266]
9: 72311 1996_Summer_Olympics
    [pem_pos=14, final_scores=0.06285156,
final_local_scores=0.046970904, pem=-4.001833,
similarity=27.254007]
10: 5843419 France
    [pem_pos=0, final_scores=0.7827837,
final_local_scores=0.7641899, pem=-0.80223274,
similarity=31.25733]
11: 3138 Atlanta
    [pem_pos=0, final_scores=0.50544435,
final_local_scores=0.51097614, pem=-0.23152597,
similarity=20.597336]
12: 5843419 France
    [pem_pos=0, final_scores=0.7198693,
final_local_scores=0.69546545, pem=-0.80223274,
similarity=29.206455]
```

Sample: b'595_French'

[1 1 French 1] shares end fractionally weaker . [2 2 PARIS 2]
1996-08-27 [3 3 French 3] shares ended fractionally weaker as
unease about union unrest slated for the autumn and a weaker
franc got the better of a slight rise on [4 4 Wall Street

4] . The blue-chip CAC-40 index ended 2.43 points or 0.12 percent lower at 2,017.99 points after a brief foray into positive territory when the [5 5 New York 5] stock market opened higher . The broader SBF-120 index closed 1.19 points or 0.08 percent lower at 1,421.90 points . Market turnover was 3.8 billion francs , about average for the quiet August period , including 2.6 billion on the most actively traded CAC-40 shares . The Socialist [6 6 CFDT 6] union warned of " tension and conflict " when [7 7 France 7] returns to work after the summer break and called for a drive to create up to 500,000 jobs in nine months . A teachers ' union , the [8 8 Fédération Syndicale Unitaire 8] (FSU) , called for members to protest against job cuts expected in the government 's austerity budget due to be unveiled in September . Anxieties over the budget niggled the currency markets where the franc lost around half a centime from Monday 's late [9 9 European 9] levels to 3.4211 per mark . Index heavyweights [10 10 Elf 10] and Rhône Poulenc both ended slightly weaker while active [11 11 Eurotunnel 11] was unchanged on nearly a million shares traded . " People are morose and it 's not the post-holiday period or the budget or company results that are going to lift anyone 's spirits , " a broker said . * [12 12 UIC 12] , part of insurer GAN , slid 12.19 percent to 55.1 francs after reporting a net attributable first-half loss of 758 million francs after the close on Monday . Markets were disappointed by a recapitalisation of 800 million francs which commentators said was larger than expected . * Supermarkets group Carrefour gained 2.19 percent to 2,616 francs after brokers Cheuvreux de Virieu confirmed the stock on their buy list , a fund manager said . * Reinsurance group Scor gained 2.1 percent to 202 francs on news that [13 13 British 13] insurer [14 14 Prudential 14] had sold its Mercantile & General reinsurance business to [15 15 Swiss 15] Re . * Conglomerate [16 16 Bollore 16] lost 2.4 percent to 521 francs after a morning trading suspension during which it said it had approved plans to buy out its 73.83 percent owned transport unit Scac Delmas Vileujeux (SDV) and invited shareholders to tender their shares . * Alcatel Alsthom fell 1.7 percent to 395.0 . * Ophthalmic products manufacturer [17 17 Essilor 17] gained 2.6 percent to 1,328 francs after Oakley Inc of the [18 18 United States 18] said it had been granted an option to buy the non-prescription lens production unit of Gentext Optics Inc , an [19 19 Essilor International 19] subsidiary . -- [20 20 Paris 20] newsroom +331 4221 5452

1: 5843419 France

2: 22989 Paris

3: 5843419 France

4: 21560 New_York_Stock_Exchange

5: 645042 New_York_City

6: 324659 French_Democratic_Confederation_of_Labour

7: 5843419 France

8: 27116031 Fédération_Syndicale_Unitaire

```
9: 9239 Europe
10: 285262 Elf_Aquitaine
11: 355334 Eurotunnel
12: 579744 International_Union_of_Railways
13: 31717 United_Kingdom
14: 1047741 Prudential_plc
15: 26748 Switzerland
16: 6052688 Bollor 
17: 3207071 Essilor
18: 3434750 United_States
19: 3207071 Essilor
20: 22989 Paris
1: 5843419 France
    [pem_pos=0, final_scores=0.71226007,
final_local_scores=0.6989055, pem=-0.80223274,
similarity=29.309114]
2: 22989 Paris
    [pem_pos=0, final_scores=0.5386267,
final_local_scores=0.5278935, pem=-0.115301244,
similarity=20.470133]
3: 5843419 France
    [pem_pos=0, final_scores=0.82297736,
final_local_scores=0.8198473, pem=-0.80223274,
similarity=32.91826]
4: 21560 New_York_Stock_Exchange
    [pem_pos=7, final_scores=0.42739898,
final_local_scores=0.48502076, pem=-2.2964394,
similarity=31.052095]
5: 645042 New_York_City
    [pem_pos=1, final_scores=0.38685805,
final_local_scores=0.39525938, pem=-1.7906687,
similarity=25.622976]
6: 324659 French_Democratic_Confederation_of_Labour
    [pem_pos=0, final_scores=0.41027725,
final_local_scores=0.50748134, pem=0.0, similarity=19.233969]
7: 5843419 France
    [pem_pos=0, final_scores=0.66840523,
final_local_scores=0.6510008, pem=-0.16240834,
similarity=24.400076]
8: 27116031 F d ration_Syndicale_Unitaire
    [pem_pos=0, final_scores=0.42642593,
final_local_scores=0.6077177, pem=0.0, similarity=22.225222]
9: 9239 Europe
    [pem_pos=0, final_scores=0.3578537,
final_local_scores=0.36192805, pem=-0.25557294,
similarity=16.280212]
10: 285262 Elf_Aquitaine
    [pem_pos=3, final_scores=0.39605564,
final_local_scores=0.46011508, pem=-1.7659115,
similarity=27.423765]
11: 355334 Eurotunnel
    [pem_pos=0, final_scores=0.25334615,
```

```
final_local_scores=0.30753437, pem=0.0, similarity=13.267152]
12: 579744 International_Union_of_Railways
    [pem_pos=0, final_scores=0.3329264,
final_local_scores=0.4379838, pem=-0.77328235,
similarity=21.365253]
13: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.48370737,
final_local_scores=0.47263622, pem=-0.48936236,
similarity=20.855349]
14: 1047741 Prudential_plc
    [pem_pos=1, final_scores=0.31508374,
final_local_scores=0.37781507, pem=-0.75650334,
similarity=19.47845]
15: 26748 Switzerland
    [pem_pos=0, final_scores=0.7043957,
final_local_scores=0.7138763, pem=-0.1392413,
similarity=26.150421]
16: 6052688 Bolloré
    [pem_pos=1, final_scores=0.3445902,
final_local_scores=0.45040864, pem=-1.0986123,
similarity=23.50523]
17: 3207071 Essilor
    [pem_pos=0, final_scores=0.5627258,
final_local_scores=0.65598154, pem=0.0, similarity=23.66551]
18: 3434750 United_States
    [pem_pos=0, final_scores=0.4985627,
final_local_scores=0.50032693, pem=-0.073457584,
similarity=19.41994]
19: 3207071 Essilor
    [pem_pos=0, final_scores=0.45582604,
final_local_scores=0.5392098, pem=0.0, similarity=20.180809]
20: 22989 Paris
    [pem_pos=0, final_scores=0.629572,
final_local_scores=0.62723744, pem=-0.115301244,
similarity=23.434755]
```

Sample: b'596_PRESS'

PRESS DIGEST - [1 1 Sri Lankan 1] Newspapers - August 27 .
Following are some of the main stories in Tuesday 's [2 2 Sri
Lankan 2] newspapers : --- [3 3 VEERAKESARI 3] Bomb blast in 1
TELO office in [4 4 Trincomalee 4] kills one , wounds six .
One officer and a soldier killed in accidental clash between
two groups of soldiers near [5 5 Chavakachcheri 5] in [6 6 Ja
ffna 6] . Army sentries thought a group of soldiers
approaching them were [7 7 Tamil 7] rebels and opened fire .
--- THINAKARAN TULF leader M. Sivasiththamparam says it is
meaningless to talk to [8 8 UNP 8] about peace package and
that the government should submit peace plan to parliament
very soon . --- [9 9 DAILY NEWS 9] Bread and flour prices
have been raised with immediate effect but government will
provide relief to underprivileged sections of society . ---
THE ISLAND Excise Commissioner W.N.F. Chandraratne denies

allegations that new guidelines in issue of liquor licences are aimed at forcing large number of liquor licence holders out of business for political reasons . --- LANKADEEPA Tamil Tiger rebels have sent 12 female suicide bombers to stage simultaneous attacks on President [10 10 Chandrika Kumaratunga 10] 's motorcade in [11 11 Colombo 11] . --- [12 12 DIVAINA 12] Cultural Ministry planning to spend large sum of money to buy silver crown believed to have been worn by ancient king and now in [13 13 Australia 13] . --- [14 14 DINAMINA 14] Government closes Ruhunu University indefinitely after big clash between two groups of students in which eight were wounded and hospitalised . -- [15 15 Colombo 15] newsroom tel 941-434319

1: 26750 Sri_Lanka
2: 26750 Sri_Lanka
3: 1904028 Virakesari
4: 566463 Trincomalee
5: 24586489 Chavakachcheri
6: 24585728 Jaffna
7: 4096004 Sri_Lankan_Tamil_people
8: 427503 United_National_Party
9: 504790 Daily_News_(New_York)
10: 307535 Chandrika_Kumaratunga
11: 56636 Colombo
12: 26165781 Divaina
13: 4689264 Australia
14: 20937676 Dinamina
15: 56636 Colombo
1: 2088911 Tamil_Eelam_Liberation_Organization
[winner: 21001313 Telo_Mascarenhas, pem_pos=1,
final_scores=-0.2405299, final_local_scores=-0.13204585,
pem=-0.6931472, similarity=3.9186437]
RECALL MISS: 2088911 Tamil_Eelam_Liberation_Organization]
1: 26750 Sri_Lanka
[pem_pos=0, final_scores=0.5139935,
final_local_scores=0.4869687, pem=-0.27403015,
similarity=20.11205]
2: 26750 Sri_Lanka
[pem_pos=0, final_scores=0.46190923,
final_local_scores=0.4300745, pem=-0.27403015,
similarity=18.414213]
3: 1904028 Virakesari
[pem_pos=0, final_scores=0.5633878,
final_local_scores=0.63080823, pem=0.0, similarity=22.91429]
4: 566463 Trincomalee
[pem_pos=0, final_scores=0.3373598,
final_local_scores=0.32563874, pem=-0.022622399,
similarity=13.930447]
5: 24586489 Chavakachcheri
[pem_pos=0, final_scores=0.4371733,
final_local_scores=0.49517578, pem=0.0, similarity=18.866745]
6: 24585728 Jaffna

```
[pem_pos=0, final_scores=0.38460857,
final_local_scores=0.3501623, pem=-0.043361604,
similarity=14.775062]
7: 4096004 Sri_Lankan_Tamil_people
    [pem_pos=4, final_scores=0.33220285,
final_local_scores=0.31436646, pem=-2.8080099,
similarity=28.741428]
8: 427503 United_National_Party
    [pem_pos=0, final_scores=0.47803372,
final_local_scores=0.49321157, pem=0.0, similarity=18.80813]
9: 504790 Daily_News_(New_York)
    [pem_pos=0, final_scores=0.26125246,
final_local_scores=0.41711205, pem=-0.3444081,
similarity=18.410114]
10: 307535 Chandrika_Kumaratunga
    [pem_pos=0, final_scores=0.64697725,
final_local_scores=0.6756404, pem=0.0, similarity=24.25217]
11: 56636 Colombo
    [pem_pos=0, final_scores=0.4841326,
final_local_scores=0.452258, pem=-0.19393882,
similarity=18.640663]
12: 26165781 Divaina
    [pem_pos=0, final_scores=0.3477381,
final_local_scores=0.45664757, pem=0.0, similarity=17.716988]
13: 4689264 Australia
    [pem_pos=0, final_scores=0.4958967,
final_local_scores=0.545894, pem=-0.17168477,
similarity=21.313929]
14: 20937676 Dinamina
    [pem_pos=0, final_scores=0.46218103,
final_local_scores=0.5427341, pem=0.0, similarity=20.28598]
15: 56636 Colombo
    [pem_pos=0, final_scores=0.52759296,
final_local_scores=0.49973178, pem=-0.19393882,
similarity=20.057375]
```

Sample: b'597_Mother'

[1 1 Mother Teresa 1] turns 86 but still in danger . Rupam Banerjee [2 2 CALCUTTA 2] 1996-08-27 [3 3 Mother Teresa 3] spent her 86th birthday in a [4 4 Calcutta 4] hospital bed on Tuesday as tributes to the legendary missionary poured in from around the world . Doctors said that later in the day they would try to wean the [5 5 Nobel Peace Prize 5] laureate from the respirator that has aided her breathing for the past six days . " Her condition seems to be better , but the danger remains as long as she is on respirator , " an official at Woodlands Nursing Home said . " She is conscious but her breathing is irregular . " The revered [6 6 Roman Catholic 6] nun was admitted to the [1 Calcutta 1] hospital a week ago with high fever and severe vomiting . She later suffered heart failure and was diagnosed with malaria . Her fever has since abated and the heart failure has been brought

under control , but her heart continues to beat irregularly , doctors said . " Unless she breathes on her own , I would advise you to keep your fingers crossed , " said a doctor who was familiar with her case but not part of the six-member team treating [7 7 Mother Teresa 7] . The nun 's birthday prompted greetings , bouquets and prayers from around the world . [8 8 Pope John Paul 8] II and [9 9 Israeli 9] Foreign Minister [10 10 David Levy 10] sent her get-well messages , the [11 11 Press Trust of India 11] said . " Ask for a miracle . Happy Birthday to our Dearest Mother , " read a placard at the Shishu Bhavan children 's home in central [12 1 2 Calcutta 12] run by [13 13 Mother Teresa 13] 's [14 14 Missionaries of Charity 14] . On Monday , both houses of [15 15 India 15] 's parliament wished the nation 's adopted sister a happy birthday and speedy recovery from her illness . Prayers continued in [16 16 Calcutta 16] , one of the world 's poorest cities , where [17 17 Mother Teresa 17] 's [18 18 Missionaries of Charity 18] runs several homes for the poor and destitute . Street children , some of them born to prostitutes , held prayers on the street . " All of us know about her . She is like a goddess , " said Raju , 8 , who has a mother but no father . [19 19 The Statesman 19] newspaper quoted 40-year-old Mangala Das , paralysed from her waist down and a resident of the Prem Dan (Gift of Love) home for the destitute , as saying she and her friends had been praying incessantly for [20 20 Mother Teresa 20] 's recovery . Tarak Das , 70 , was picked up from a [21 21 Calcutta 21] footpath a week ago by passers-by who took pity on him and brought him to Nirmal Hriday (Immaculate Home) . " I do not know who she is . I have never seen her , but I can only bless her for what she has done for people like me , " Das told [22 22 The Statesman 22] . [23 23 Mother Teresa 23] 's condition improved on Sunday as her fever abated , and on Monday she was able to scribble notes to doctors and nuns . Thousands in [24 24 Calcutta 24] , where she founded her [25 2 5 Missionaries of Charity 25] religious order in 1949 , prayed for her recovery . Ministers of the communist government of [26 26 West Bengal 26] state and people of different religions joined [27 27 Catholics 27] to pray for [28 28 Mother Teresa 28] 's recovery at Mother House . " We joined the prayer to express our solidarity with her work for the cause of the poor and downtrodden , " said Nanda Gopal Bhattacharya , a communist minister in [29 29 West Bengal 29] .

- 1: 347105 Mother_Teresa
- 2: 47905 Kolkata
- 3: 347105 Mother_Teresa
- 4: 47905 Kolkata
- 5: 26230922 Nobel_Peace_Prize
- 6: 606848 Catholic_Church
- 7: 347105 Mother_Teresa
- 8: 23805 Pope_John_Paul_II

```
9: 9282173 Israel
10: 1256116 David_Levy_(Israeli_politician)
11: 1724500 Press_Trust_of_India
12: 47905 Kolkata
13: 347105 Mother_Teresa
14: 345921 Missionaries_of_Charity
15: 14533 India
16: 47905 Kolkata
17: 347105 Mother_Teresa
18: 345921 Missionaries_of_Charity
19: 1150534 The_Statesman
20: 347105 Mother_Teresa
21: 47905 Kolkata
22: 1150534 The_Statesman
23: 347105 Mother_Teresa
24: 47905 Kolkata
25: 345921 Missionaries_of_Charity
26: 34040 West_Bengal
27: 606848 Catholic_Church
28: 347105 Mother_Teresa
29: 34040 West_Bengal
1: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.594289,
final_local_scores=0.57431835, pem=0.0, similarity=21.228518]
2: 47905 Kolkata
    [pem_pos=0, final_scores=0.50290954,
final_local_scores=0.47325712, pem=-0.02273471,
similarity=18.336285]
3: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.6712437,
final_local_scores=0.65837973, pem=0.0, similarity=23.737078]
4: 47905 Kolkata
    [pem_pos=0, final_scores=0.39632648,
final_local_scores=0.35683134, pem=-0.02273471,
similarity=14.861906]
5: 26230922 Nobel_Peace_Prize
    [pem_pos=0, final_scores=0.29631764,
final_local_scores=0.33097413, pem=0.0, similarity=13.966641]
6: 606848 Catholic_Church
    [pem_pos=0, final_scores=0.41853088,
final_local_scores=0.4413699, pem=0.0, similarity=17.26107]
7: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.5989228,
final_local_scores=0.5793801, pem=0.0, similarity=21.37957]
8: 23805 Pope_John_Paul_II
    [pem_pos=0, final_scores=0.32059246,
final_local_scores=0.33009425, pem=0.0, similarity=13.940384]
9: 9282173 Israel
    [pem_pos=0, final_scores=0.6071784,
final_local_scores=0.68450993, pem=-0.34790626,
similarity=26.408823]
10: 1256116 David_Levy_(Israeli_politician)
```

```
[pem_pos=1, final_scores=0.5539252,
final_local_scores=0.70041627, pem=-1.1434275,
similarity=31.209667]
11: 1724500 Press_Trust_of_India
    [pem_pos=0, final_scores=0.5331922,
final_local_scores=0.61271673, pem=0.0, similarity=22.374403]
12: 47905 Kolkata
    [pem_pos=0, final_scores=0.46663827,
final_local_scores=0.43363625, pem=-0.02273471,
similarity=17.15392]
13: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.58323354,
final_local_scores=0.562242, pem=0.0, similarity=20.868135]
14: 345921 Missionaries_of_Charity
    [pem_pos=0, final_scores=0.5609115,
final_local_scores=0.5904018, pem=0.0, similarity=21.70848]
15: 14533 India
    [pem_pos=0, final_scores=0.6560443,
final_local_scores=0.6790351, pem=-0.13622136,
similarity=25.094269]
16: 47905 Kolkata
    [pem_pos=0, final_scores=0.43990892,
final_local_scores=0.40443856, pem=-0.02273471,
similarity=16.282602]
17: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.62909454,
final_local_scores=0.6123382, pem=0.0, similarity=22.363106]
18: 345921 Missionaries_of_Charity
    [pem_pos=0, final_scores=0.52230006,
final_local_scores=0.5482247, pem=0.0, similarity=20.449831]
19: 1150534 The_Statesman
    [pem_pos=0, final_scores=0.48482317,
final_local_scores=0.5232172, pem=0.0, similarity=19.703558]
20: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.6051869,
final_local_scores=0.58622265, pem=0.0, similarity=21.583767]
21: 47905 Kolkata
    [pem_pos=0, final_scores=0.37424308,
final_local_scores=0.3327086, pem=-0.02273471,
similarity=14.1420355]
22: 1150534 The_Statesman
    [pem_pos=0, final_scores=0.5209573,
final_local_scores=0.56268823, pem=0.0, similarity=20.881453]
23: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.682753,
final_local_scores=0.67095184, pem=0.0, similarity=24.112255]
24: 47905 Kolkata
    [pem_pos=0, final_scores=0.5012786,
final_local_scores=0.47147554, pem=-0.02273471,
similarity=18.28312]
25: 345921 Missionaries_of_Charity
    [pem_pos=0, final_scores=0.6225819,
```

```
final_local_scores=0.65776736, pem=0.0, similarity=23.718803]
26: 34040 West_Bengal
    [pem_pos=0, final_scores=0.42954296,
final_local_scores=0.449179, pem=0.0, similarity=17.49411]
27: 606848 Catholic_Church
    [pem_pos=0, final_scores=0.23068492,
final_local_scores=0.23617676, pem=-0.38803694,
similarity=13.247904]
28: 347105 Mother_Teresa
    [pem_pos=0, final_scores=0.5940034,
final_local_scores=0.5740063, pem=0.0, similarity=21.219206]
29: 34040 West_Bengal
    [pem_pos=0, final_scores=0.47262412,
final_local_scores=0.4962386, pem=0.0, similarity=18.898462]
1: 47905 Kolkata
    [pem_pos=0, final_scores=0.13444759,
final_local_scores=0.070768416, pem=-0.02273471,
similarity=6.3252187]
```

Sample: b'598_Islamists'

[1 1 Islamists 1] can meet in [2 2 London 2] , minister . [3 3 ISLAMABAD 3] 1996-08-27 [4 4 British 4] Foreign Secretary [5 5 Malcolm Rifkind 5] said on Tuesday that his government would only take action against a planned conference of [6 6 Islamist 6] groups in [7 7 London 7] if [8 8 British 8] law was broken . " People who wish to hold conferences of course do n't need to seek permission from the government in [9 9 Britain 9] , " [10 10 Rifkind 10] , in [11 11 Pakistan 11] for a visit , told [12 12 Reuters 12] . " As long as they obey our laws then that is not something the government would normally interfere with . " The [13 13 Islamist 13] conference , due to be held in [14 14 London 14] on September 8 , has caused concern in countries such as [15 15 Algeria 15] and [16 16 Egypt 16] , which are fighting armed [17 17 Islamic 17] militants . [18 18 British 18] [19 19 Jewish 19] groups have also protested because they say members of [20 20 Algeria 20] 's [21 21 Islamic Salvation Front 21] ([22 22 FIS 22]) and the [23 23 Palestinian 23] [24 24 Islamic 24] group [25 25 Hamas 25] are on the guest list . [26 26 Rifkind 26] said it was for the home secretary (interior minister) to act by denying visas to participants if he felt there was reason to believe that they might break the law . " Our policy has to be fundamentally based on respect for the rule of law and insistence that it be observed , " he said . [27 27 Rifkind 27] was in [28 28 Pakistan 28] at the start of an [29 29 Asian 29] tour that will also take him to [30 30 India 30] , [31 31 Sri Lanka 31] , [32 32 Japan 32] and [33 33 Mongolia 33] .

1: 15012 Islamism
2: 17867 London
3: 51153 Islamabad
4: 31717 United_Kingdom
5: 354551 Malcolm_Rifkind

6: 15012 Islamism
7: 17867 London
8: 31717 United_Kingdom
9: 31717 United_Kingdom
10: 354551 Malcolm_Rifkind
11: 23235 Pakistan
12: 18998750 Reuters
13: 15012 Islamism
14: 17867 London
15: 358 Algeria
16: 8087628 Egypt
17: 15012 Islamism
18: 31717 United_Kingdom
19: 25955086 Jews
20: 358 Algeria
21: 403399 Islamic_Salvation_Front
22: 403399 Islamic_Salvation_Front
23: 241405 State_of_Palestine
24: 6037917 Islam
25: 13913 Hamas
26: 354551 Malcolm_Rifkind
27: 354551 Malcolm_Rifkind
28: 23235 Pakistan
29: 689 Asia
30: 14533 India
31: 26750 Sri_Lanka
32: 15573 Japan
33: 19271 Mongolia

1: 15012 Islamism
[pem_pos=0, final_scores=0.40771097,
final_local_scores=0.40929323, pem=0.0, similarity=16.303839]
2: 17867 London
[pem_pos=0, final_scores=0.496185,
final_local_scores=0.49264818, pem=-0.09086737,
similarity=19.285467]
3: 51153 Islamabad
[pem_pos=0, final_scores=0.3724466,
final_local_scores=0.3790229, pem=0.0, similarity=15.400513]
4: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5883531,
final_local_scores=0.56284106, pem=-0.48936236,
similarity=23.547241]
5: 354551 Malcolm_Rifkind
[pem_pos=0, final_scores=0.5860366,
final_local_scores=0.6617155, pem=0.0, similarity=23.836624]
6: 15012 Islamism
[pem_pos=0, final_scores=0.36006278,
final_local_scores=0.35724482, pem=0.0, similarity=14.750611]
7: 17867 London
[pem_pos=0, final_scores=0.5690461,
final_local_scores=0.5722378, pem=-0.09086737,
similarity=21.66058]

```
8: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5383252,
final_local_scores=0.5081932, pem=-0.48936236,
similarity=21.916441]
9: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5880663,
final_local_scores=0.56252784, pem=-0.61114657,
similarity=24.200176]
10: 354551 Malcolm_Rifkind
    [pem_pos=0, final_scores=0.439681,
final_local_scores=0.5018443, pem=0.0, similarity=19.065748]
11: 23235 Pakistan
    [pem_pos=0, final_scores=0.59782714,
final_local_scores=0.59069633, pem=-0.21944217,
similarity=22.91063]
12: 18998750 Reuters
    [pem_pos=0, final_scores=0.4350831,
final_local_scores=0.44238186, pem=0.0, similarity=17.29127]
13: 15012 Islamism
    [pem_pos=0, final_scores=0.33716977,
final_local_scores=0.33223772, pem=0.0, similarity=14.004349]
14: 17867 London
    [pem_pos=0, final_scores=0.5304368,
final_local_scores=0.5300631, pem=-0.09086737,
similarity=20.402002]
15: 358 Algeria
    [pem_pos=0, final_scores=0.5862269,
final_local_scores=0.60669774, pem=-0.33694586,
similarity=24.027147]
16: 8087628 Egypt
    [pem_pos=0, final_scores=0.5596066,
final_local_scores=0.5402535, pem=-0.27480945,
similarity=21.706413]
17: 15012 Islamism
    [pem_pos=3, final_scores=0.18745597,
final_local_scores=0.17117015, pem=-2.7089531,
similarity=23.92948]
18: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.42018837,
final_local_scores=0.37914664, pem=-0.48936236,
similarity=18.065434]
19: 25955086 Jews
    [pem_pos=1, final_scores=0.19051073,
final_local_scores=0.17894882, pem=-0.80476934,
similarity=13.806363]
20: 358 Algeria
    [pem_pos=0, final_scores=0.63417035,
final_local_scores=0.6590687, pem=-0.33694586,
similarity=25.590002]
21: 403399 Islamic_Salvation_Front
    [pem_pos=0, final_scores=0.58397114,
final_local_scores=0.66596115, pem=0.0, similarity=23.963322]
```

```
22: 403399 Islamic_Salvation_Front
    [pem_pos=2, final_scores=0.4611538,
final_local_scores=0.5318019, pem=-1.4075395,
similarity=27.61416]
23: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.63605934,
final_local_scores=0.6759659, pem=-2.8508017,
similarity=39.764988]
24: 6037917 Islam
    [pem_pos=0, final_scores=0.41241282,
final_local_scores=0.4201861, pem=-0.21141872,
similarity=17.778633]
25: 13913 Hamas
    [pem_pos=0, final_scores=0.467314,
final_local_scores=0.49462253, pem=0.0, similarity=18.850235]
26: 354551 Malcolm_Rifkind
    [pem_pos=0, final_scores=0.35904807,
final_local_scores=0.41376507, pem=0.0, similarity=16.437288]
27: 354551 Malcolm_Rifkind
    [pem_pos=0, final_scores=0.4278366,
final_local_scores=0.4889061, pem=0.0, similarity=18.679646]
28: 23235 Pakistan
    [pem_pos=0, final_scores=0.574076,
final_local_scores=0.5647518, pem=-0.21944217,
similarity=22.136393]
29: 689 Asia
    [pem_pos=1, final_scores=0.36812705,
final_local_scores=0.38547665, pem=-1.9623682,
similarity=26.264769]
30: 14533 India
    [pem_pos=0, final_scores=0.56488544,
final_local_scores=0.57800204, pem=-0.13622136,
similarity=22.079239]
31: 26750 Sri_Lanka
    [pem_pos=0, final_scores=0.5360533,
final_local_scores=0.5555088, pem=-0.3027737,
similarity=22.313734]
32: 15573 Japan
    [pem_pos=0, final_scores=0.6246641,
final_local_scores=0.6684816, pem=-0.13549899,
similarity=24.7754]
33: 19271 Mongolia
    [pem_pos=0, final_scores=0.66591614,
final_local_scores=0.7154753, pem=-0.36161825,
similarity=27.407461]
```

Sample: b'599_Afghan'

[1 1 Afghan 1] leader tells [2 2 U.S. 2] Congressman of
peace plan . Sayed Salahuddin [3 3 KABUL 3] 1996-08-27 [4 4 Af
ghan 4] government military chief [5 5 Ahmad Shah Masood 5]
briefed visiting [6 6 U.S. 6] Congressman [7 7 Dana Rohrabach
er 7] on Tuesday on a peace plan for his wartorn country . A

spokesman for Masood said he had told the [8 8 California 8] [9 9 Republican 9] at a meeting in northern [10 10 Kabul 10] that President [11 11 Burhanuddin Rabbani 11] 's government favoured talks with all [12 12 Afghan 12] factions to set up an interim government . The factions should agree to appoint a transitional leader , draft a new constitution , collect heavy weapons , create a national army and hold free elections in which the transitional leader would be barred from standing , he added . [13 13 Rohrabacher 13] flew into [14 14 Bagram 14] military airbase north of [15 15 Kabul 15] in a [16 16 Red Cross 16] plane on Tuesday after meeting northern opposition militia leader General [17 17 Abdul Rashid Dostum 17] . Masood 's spokesman Amrollah (one name) said [18 18 Rohrabacher 18] had recently visited [19 19 Italy 19] , [20 20 Saudi Arabia 20] and [21 21 Pakistan 21] as part of a mission to promote peace in [22 22 Afghanistan 22] . " We are certainly serious more than before to find a solution to the [23 23 Afghan 23] problem and support every [24 24 U.N. 24] plan , " Amrollah quoted [25 25 Rohrabacher 25] as saying . However , a spokesman for Prime Minister [26 26 Gulbuddin Hekmatyar 26] , a long-time rival of Masood , expressed concern at signs of renewed [27 27 U.S. 27] interest in [28 28 Afghanistan 28] . " [29 29 America 29] wants to block the establishment of a strong [30 30 Islamic 30] government in [31 31 Afghanistan 31] and the [32 32 U.S. 32] intends to neutralise the [33 33 Afghan 33] peace process initiated by the [34 34 Afghans 34] themselves , " said the spokesman , Hamid Ibrahimi . " A great game has been started in [35 35 Afghanistan 35] as [36 36 America 36] feels that [37 37 Tehran 37] and 1 Moscow have got stronger in the [38 38 Afghan 38] picture -- something 2 Washington wants to change , " he said . [39 39 Rohrabacher 39] was expected to visit neutral faction leaders in the eastern city of [40 40 Jalalabad 40] and meet leaders of the rebel Islamic Taliban militia in the southern city of [41 41 Kandahar 41] . [42 42 Afghan 42] guerrilla factions have been locked in a bloody power struggle since the fall of the communist government in April 1992 . [43 43 Hekmatyar 43] , once [44 44 Rabbani 44] 's main rival , made a peace pact with him and rejoined the government as prime minister in June .

1: 737 Afghanistan

2: 3434750 United_States

3: 16826 Kabul

4: 737 Afghanistan

5: 2176 Ahmad_Shah_Massoud

6: 3434750 United_States

7: 8953 Dana_Rohrabacher

8: 5407 California

9: 32070 Republican_Party_(United_States)

10: 16826 Kabul

11: 4611 Burhanuddin_Rabbani

12: 737 Afghanistan

13: 8953 Dana_Rohrabacher
14: 407212 Bagram
15: 16826 Kabul
16: 5023832 International_Committee_of_the_Red_Cross
17: 2376 Abdul_Rashid_Dostum
18: 8953 Dana_Rohrabacher
19: 14532 Italy
20: 349303 Saudi_Arabia
21: 23235 Pakistan
22: 737 Afghanistan
23: 737 Afghanistan
24: 31769 United_Nations
25: 8953 Dana_Rohrabacher
26: 161917 Gulbuddin_Hekmatyar
27: 3434750 United_States
28: 737 Afghanistan
29: 3434750 United_States
30: 15012 Islamism
31: 737 Afghanistan
32: 3434750 United_States
33: 737 Afghanistan
34: 737 Afghanistan
35: 737 Afghanistan
36: 3434750 United_States
37: 14653 Iran
38: 737 Afghanistan
39: 8953 Dana_Rohrabacher
40: 16274 Jalalabad
41: 17260 Kandahar
42: 737 Afghanistan
43: 161917 Gulbuddin_Hekmatyar
44: 4611 Burhanuddin_Rabbani

1: 25391 Russia
[winner: 19004 Moscow, pem_pos=0,
final_scores=-0.4246568, final_local_scores=-0.46381444,
pem=-0.106582455, similarity=-0.1718235]
RECALL MISS: 25391 Russia]

2: 3434750 United_States
[winner: 108956 Washington,_D.C., pem_pos=0,
final_scores=-0.24797489, final_local_scores=-0.2941124,
pem=-1.7496481, similarity=4.8276696]
RECALL MISS: 3434750 United_States]

1: 737 Afghanistan
[pem_pos=0, final_scores=0.5714461,
final_local_scores=0.5227962, pem=-0.869566,
similarity=24.41983]

2: 3434750 United_States
[pem_pos=0, final_scores=0.46647316,
final_local_scores=0.4669748, pem=-0.14075686,
similarity=18.790628]

3: 16826 Kabul
[pem_pos=0, final_scores=0.3700993,

```
final_local_scores=0.32210553, pem=0.0, similarity=13.701984]
4: 737 Afghanistan
    [pem_pos=0, final_scores=0.57132196,
final_local_scores=0.52266055, pem=-0.869566,
similarity=24.415783]
5: 2176 Ahmad_Shah_Massoud
    [pem_pos=0, final_scores=0.5576424,
final_local_scores=0.57866293, pem=0.0, similarity=21.35817]
6: 3434750 United_States
    [pem_pos=0, final_scores=0.58643585,
final_local_scores=0.5980157, pem=-0.14075686,
similarity=22.70115]
7: 8953 Dana_Rohrabacher
    [pem_pos=0, final_scores=0.6895202,
final_local_scores=0.77455825, pem=0.0, similarity=27.204077]
8: 5407 California
    [pem_pos=0, final_scores=0.6364706,
final_local_scores=0.6979542, pem=-0.08662026,
similarity=25.389112]
9: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.33921725,
final_local_scores=0.39934683, pem=-0.1310191,
similarity=16.71952]
10: 16826 Kabul
    [pem_pos=0, final_scores=0.42890602,
final_local_scores=0.38634294, pem=0.0, similarity=15.6189575]
11: 4611 Burhanuddin_Rabbani
    [pem_pos=0, final_scores=0.5475262,
final_local_scores=0.5677472, pem=0.0, similarity=21.032421]
12: 737 Afghanistan
    [pem_pos=0, final_scores=0.65284294,
final_local_scores=0.6117098, pem=-0.869566,
similarity=27.073189]
13: 8953 Dana_Rohrabacher
    [pem_pos=0, final_scores=0.49020827,
final_local_scores=0.5568403, pem=0.0, similarity=20.706938]
14: 407212 Bagram
    [pem_pos=0, final_scores=0.27851993,
final_local_scores=0.30669212, pem=-0.30189708,
similarity=14.883781]
15: 16826 Kabul
    [pem_pos=0, final_scores=0.41837054,
final_local_scores=0.37483448, pem=0.0, similarity=15.275521]
16: 5023832 International_Committee_of_the_Red_Cross
    [pem_pos=1, final_scores=0.3211276,
final_local_scores=0.34586293, pem=-1.9555908,
similarity=25.04576]
17: 2376 Abdul_Rashid_Dostum
    [pem_pos=0, final_scores=0.5803898,
final_local_scores=0.62369907, pem=0.0, similarity=22.702137]
18: 8953 Dana_Rohrabacher
    [pem_pos=0, final_scores=0.4898842,
```

```
final_local_scores=0.55648625, pem=0.0, similarity=20.696373]
19: 14532 Italy
    [pem_pos=0, final_scores=0.58039504,
final_local_scores=0.6421244, pem=-0.20612453,
similarity=24.372925]
20: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.45854336,
final_local_scores=0.43571216, pem=-0.25454172,
similarity=18.47647]
21: 23235 Pakistan
    [pem_pos=0, final_scores=0.61968416,
final_local_scores=0.6022352, pem=-0.21944217,
similarity=23.254972]
22: 737 Afghanistan
    [pem_pos=0, final_scores=0.59833175,
final_local_scores=0.5521646, pem=-0.22286008,
similarity=21.779354]
23: 737 Afghanistan
    [pem_pos=0, final_scores=0.57922614,
final_local_scores=0.53129464, pem=-0.869566,
similarity=24.67344]
24: 31769 United_Nations
    [pem_pos=0, final_scores=0.48683912,
final_local_scores=0.5007065, pem=0.0, similarity=19.031794]
25: 8953 Dana_Rohrabacher
    [pem_pos=0, final_scores=0.5527929,
final_local_scores=0.6252045, pem=0.0, similarity=22.747063]
26: 161917 Gulbuddin_Hekmatyar
    [pem_pos=0, final_scores=0.47278243,
final_local_scores=0.48692948, pem=0.0, similarity=18.620659]
27: 3434750 United_States
    [pem_pos=0, final_scores=0.46744853,
final_local_scores=0.46804017, pem=-0.14075686,
similarity=18.822424]
28: 737 Afghanistan
    [pem_pos=0, final_scores=0.65637225,
final_local_scores=0.6155651, pem=-0.22286008,
similarity=23.671352]
29: 3434750 United_States
    [pem_pos=0, final_scores=0.45728952,
final_local_scores=0.45694298, pem=-0.78483206,
similarity=21.993841]
30: 15012 Islamism
    [pem_pos=3, final_scores=0.19946302,
final_local_scores=0.19127038, pem=-2.7089531,
similarity=24.52931]
31: 737 Afghanistan
    [pem_pos=0, final_scores=0.5662474,
final_local_scores=0.5171174, pem=-0.22286008,
similarity=20.733475]
32: 3434750 United_States
    [pem_pos=0, final_scores=0.51012963,
```

```
final_local_scores=0.5146628, pem=-0.14075686,
similarity=20.213736]
33: 737 Afghanistan
    [pem_pos=0, final_scores=0.58312875,
final_local_scores=0.5355577, pem=-0.869566,
similarity=24.800657]
34: 737 Afghanistan
    [pem_pos=1, final_scores=0.47309864,
final_local_scores=0.41536647, pem=-1.2683228,
similarity=23.382412]
35: 737 Afghanistan
    [pem_pos=0, final_scores=0.5725613,
final_local_scores=0.52401435, pem=-0.22286008,
similarity=20.939295]
36: 3434750 United_States
    [pem_pos=0, final_scores=0.5060217,
final_local_scores=0.5101755, pem=-0.78483206,
similarity=23.582407]
37: 14653 Iran
    [pem_pos=5, final_scores=0.3872246,
final_local_scores=0.42974144, pem=-2.6998966,
similarity=31.59651]
38: 737 Afghanistan
    [pem_pos=0, final_scores=0.6162049,
final_local_scores=0.57168835, pem=-0.869566,
similarity=25.878868]
39: 8953 Dana_Rohrabacher
    [pem_pos=0, final_scores=0.56490916,
final_local_scores=0.63843966, pem=0.0, similarity=23.142025]
40: 16274 Jalalabad
    [pem_pos=0, final_scores=0.27831602,
final_local_scores=0.2581878, pem=0.0, similarity=11.794552]
41: 17260 Kandahar
    [pem_pos=0, final_scores=0.32405573,
final_local_scores=0.29563487, pem=-0.11192558,
similarity=13.520716]
42: 737 Afghanistan
    [pem_pos=0, final_scores=0.5267433,
final_local_scores=0.4739651, pem=-0.869566,
similarity=22.962612]
43: 161917 Gulbuddin_Hekmatyar
    [pem_pos=0, final_scores=0.40350157,
final_local_scores=0.4112506, pem=0.0, similarity=16.362251]
44: 4611 Burhanuddin_Rabbani
    [pem_pos=0, final_scores=0.4008022,
final_local_scores=0.41237336, pem=0.0, similarity=16.395756]
```

Sample: b'600_Pakistan'

```
[1 1 Pakistan 1] state bank sells 1.38 bln rupees of bonds .
[2 2 KARACHI 2] , [3 3 Pakistan 3] 1996-08-27 The State
( central ) Bank of [4 4 Pakistan 4] auctioned three- , five-
and 10-year federal investment bonds worth 1.38 billion
```

rupees on Tuesday . The bank said it had accepted bids of 250 million rupees at par for three-year bonds , 3.5 million rupees at par for five-year bonds and 1.126 billion at par for 10-year bonds . The auction is set for settlement on Thursday . In the previous auction on July 11 , it accepted bids worth 300 million rupees at par for three-year bonds , 44.5 million rupees at par for five-year bonds and 782.6 million rupees at par for 10-year bonds . -- [5 5 Karachi 5] newsroom 9221-5685192

1: 23235 Pakistan
2: 17123 Karachi
3: 23235 Pakistan
4: 23235 Pakistan
5: 17123 Karachi
1: 23235 Pakistan
[pem_pos=0, final_scores=0.593246,
final_local_scores=0.535102, pem=-0.21944217,
similarity=21.251585]
2: 17123 Karachi
[pem_pos=0, final_scores=0.46183985,
final_local_scores=0.41067642, pem=0.0, similarity=16.345116]
3: 23235 Pakistan
[pem_pos=0, final_scores=0.6605931,
final_local_scores=0.6086685, pem=-0.21944217,
similarity=23.446955]
4: 23235 Pakistan
[pem_pos=0, final_scores=0.6769314,
final_local_scores=0.62651557, pem=-0.21944217,
similarity=23.979548]
5: 17123 Karachi
[pem_pos=0, final_scores=0.4587018,
final_local_scores=0.40724862, pem=0.0, similarity=16.242825]

Sample: b'601_Nepal'

[1 1 Nepal 1] offers to talk to [2 2 Maoist 2] insurgents . Gopal Sharma [3 3 KATHMANDU 3] 1996-08-27 [4 4 Nepal 4] 's centre-right coalition government has offered to meet the country 's hardline [5 5 Maoist 5] communists for talks in a bid to end an insurgency in [6 6 Nepal 6] 's western districts , officials said on Tuesday . The [7 7 Maoists 7] oppose multi-party democracy and want to establish a communist state . But the [8 8 Nepali 8] government said the insurgents must give up violence before it negotiates with them . " They (the insurgents) should first give up their violent activities , " Home (Interior) Minister Khum Bahadur Khadga said . About 54 people have died in [9 9 Maoist 9] insurgent activity and in police action against them since February , officials said . [10 10 Nepali 10] opposition parties have accused the police of having killed more people than the insurgents . Some human rights groups have criticised the government 's handling of the situation . In a speech in parliament on Tuesday , Khadga challenged the [11 11]

Maoist 11] communists to " win the people 's confidence " and win election to parliament . On Monday , he had offered to talk to leaders of the United People 's Front Nepal ([12 12 B hattarai 12]) , the [13 13 Maoist 13] faction which leads the insurgency . " The government is ready to guarantee security of the [14 14 Maoist 14] representatives who want to take part in peaceful dialogue , " Khadga said . A multi-party democracy was set up in [15 15 Nepal 15] six years ago , after a popular movement by the centrist [16 16 Nepali C ongress 16] party jointly with the Communist United Marxist-Leninist (UML) party . The [17 17 Nepali Congress 17] leads the three-party coalition government while the UML is the main opposition party .

1: 171166 Nepal
2: 403685 Unified_Communist_Party_of_Nepal_(Maoist)
3: 17168 Kathmandu
4: 171166 Nepal
5: 403685 Unified_Communist_Party_of_Nepal_(Maoist)
6: 171166 Nepal
7: 403685 Unified_Communist_Party_of_Nepal_(Maoist)
8: 171166 Nepal
9: 403685 Unified_Communist_Party_of_Nepal_(Maoist)
10: 171166 Nepal
11: 403685 Unified_Communist_Party_of_Nepal_(Maoist)
12: 392573 Baburam_Bhattarai
13: 403685 Unified_Communist_Party_of_Nepal_(Maoist)
14: 403685 Unified_Communist_Party_of_Nepal_(Maoist)
15: 171166 Nepal
16: 364670 Nepali_Congress
17: 364670 Nepali_Congress
1: 171166 Nepal
[pem_pos=0, final_scores=0.6191646,
final_local_scores=0.6131594, pem=-0.22887385,
similarity=23.632265]
2: 403685 Unified_Communist_Party_of_Nepal_(Maoist)
[pem_pos=4, final_scores=0.5383873,
final_local_scores=0.5009351, pem=-1.3851347,
similarity=26.571192]
3: 17168 Kathmandu
[pem_pos=0, final_scores=0.44116038,
final_local_scores=0.4383508, pem=0.0, similarity=17.170975]
4: 171166 Nepal
[pem_pos=0, final_scores=0.7772153,
final_local_scores=0.7858058, pem=-0.22887385,
similarity=28.784378]
5: 403685 Unified_Communist_Party_of_Nepal_(Maoist)
[pem_pos=4, final_scores=0.52223366,
final_local_scores=0.48328966, pem=-1.3851347,
similarity=26.044619]
6: 171166 Nepal
[pem_pos=0, final_scores=0.6118937,
final_local_scores=0.6052171, pem=-0.22887385,

```
similarity=23.39525]
7: 403685 Unified_Communist_Party_of_Nepal_(Maoist)
    [pem_pos=3, final_scores=0.50419056,
final_local_scores=0.4635803, pem=-1.0941973,
similarity=23.874289]
8: 171166 Nepal
    [pem_pos=1, final_scores=0.52717453,
final_local_scores=0.51267433, pem=-1.4646683,
similarity=27.35403]
9: 403685 Unified_Communist_Party_of_Nepal_(Maoist)
    [pem_pos=4, final_scores=0.44777256,
final_local_scores=0.40195227, pem=-1.3851347,
similarity=23.617348]
10: 171166 Nepal
    [pem_pos=1, final_scores=0.4173476,
final_local_scores=0.39270508, pem=-1.4646683,
similarity=23.773909]
11: 403685 Unified_Communist_Party_of_Nepal_(Maoist)
    [pem_pos=4, final_scores=0.5338158,
final_local_scores=0.4959414, pem=-1.3851347,
similarity=26.42217]
12: 392573 Baburam_Bhattarai
    [pem_pos=0, final_scores=0.4184212,
final_local_scores=0.44249272, pem=0.0, similarity=17.294579]
13: 403685 Unified_Communist_Party_of_Nepal_(Maoist)
    [pem_pos=4, final_scores=0.45507592,
final_local_scores=0.40993, pem=-1.3851347,
similarity=23.85542]
14: 403685 Unified_Communist_Party_of_Nepal_(Maoist)
    [pem_pos=4, final_scores=0.48913103,
final_local_scores=0.44713008, pem=-1.3851347,
similarity=24.965544]
15: 171166 Nepal
    [pem_pos=0, final_scores=0.5946691,
final_local_scores=0.5864018, pem=-0.22887385,
similarity=22.833765]
16: 364670 Nepali_Congress
    [pem_pos=0, final_scores=0.3177395,
final_local_scores=0.298121, pem=0.0, similarity=12.9862385]
17: 364670 Nepali_Congress
    [pem_pos=0, final_scores=0.3183927,
final_local_scores=0.29883456, pem=0.0, similarity=13.007532]
```

Sample: b'602_Indian'

[1 1 Indian 1] soy prices end steady ahead of holiday . [2 2
INDORE 2] , [3 3 India 3] 1996-08-27 [4 4 Indian 4] soybean
prices on Tuesday remained steady at 12,900-13,100 rupees per
tonne in plant delivery condition , dealers said They said
arrivals were poor due to the festival season . Markets in
central [5 5 India 5] would be closed for a local religious
holiday on Wednesday . Soyoil prices fell on increased
selling against poor demand . Soyoil solvent was down by 400

rupees per tonne and soyoil refined was down by 400 rupees . Soyoil refined fell by 200 rupees on weak undertone . Soymeal yellow was \$ 276-277 and soymeal black was \$ 246-248 per tonne in export . Rapeseed extraction was \$ 115 per tonne in export . Export demand was good but availability was limited . Rapeseed extraction was 3,850 rupees FOR Bedibunder and was 3,800-3,825 rupees FOR [6 6 Bhavnagar 6] .

----- (Prices in rupees per tonne) Market Arrivals Auction Traders Plant (in tonnes) [7 7 Dewas 7] 45 Yellow 12,700-12,950 12,900-13,150 12,900-13,100 Black 11,900-12,100 [8 8 Mandsaur 8] 10 Yellow 12,600-12,750 12,700-12,850 [9 9 Neemuch 9] n.a Yellow - - [10 10 Mhow 10] 2 Yellow 12,700-12,800 12,750-12,850 [11 11 Ratlam 11] 10 Yellow 12,600-12,750 12,700-12,800 [12 12 Ashta 12] 10 Yellow 12,700-12,900 12,800-13,000 [13 13 Indore 13] 25 Yellow 12,750-12,950 12,900-13,100 [14 14 Dhar 14] 5 Yellow 12,700-12,800 12,750-12,900 [15 15 Ujjain 15] 8 Yellow 12,750-12,900 12,850-13,050 [16 16 Jaora 16] n.a Yellow - - Barnagar n.a Yellow - - [17 17 Khandwa 17] n.a Yellow - - [18 18 Ashoknagar 18] n.a Yellow - - Nalkhera n.a Yellow - -

----- Soyoil (in rupees per tonne) Soyoil solvent plant delivery 30,300-30,400 Soyoil solvent market delivery 30,700-30,800 Soyoil refined plant delivery 32,700-32,800 Soyoil refined market delivery 32,900-33,000 ----- Soymeal (in rupees per tonne , free on rail-FOR) Yellow Black FOR [19 19 Bombay 19] 9,800 8,800 FOR Bedi Bunder 9,800 8,800 (\$ 1=35.73 rupees)

1: 14533 India
2: 365670 Indore
3: 14533 India
4: 14533 India
5: 14533 India
6: 1278576 Bhavnagar
7: 906379 Dewas
8: 954847 Mandsaur
9: 1923904 Neemuch
10: 740431 Mhow
11: 954785 Ratlam
12: 5706727 Ashta,_Madhya_Pradesh
13: 365670 Indore
14: 911235 Dhar
15: 901117 Ujjain
16: 2001103 Jaora
17: 509811 Khandwa
18: 2293572 Ashoknagar
19: 19189 Mumbai
1: 14533 India
[pem_pos=0, final_scores=0.45620388, final_local_scores=0.43559724, pem=-0.618294, similarity=20.451183]
2: 365670 Indore

```
[pem_pos=0, final_scores=0.41689843,
final_local_scores=0.37451082, pem=-0.031899266,
similarity=15.439336]
3: 14533 India
    [pem_pos=0, final_scores=0.692064,
final_local_scores=0.6932386, pem=-0.13622136,
similarity=25.51813]
4: 14533 India
    [pem_pos=0, final_scores=0.4799719,
final_local_scores=0.4615602, pem=-0.618294,
similarity=21.22597]
5: 14533 India
    [pem_pos=0, final_scores=0.5431488,
final_local_scores=0.53057134, pem=-0.13622136,
similarity=20.663813]
6: 1278576 Bhavnagar
    [pem_pos=0, final_scores=0.33193564,
final_local_scores=0.3142356, pem=0.0, similarity=13.46713]
7: 906379 Dewas
    [pem_pos=0, final_scores=0.49397713,
final_local_scores=0.4924994, pem=0.0, similarity=18.786879]
8: 954847 Mandsaur
    [pem_pos=0, final_scores=0.39725226,
final_local_scores=0.40036124, pem=0.0, similarity=16.03729]
9: 1923904 Neemuch
    [pem_pos=0, final_scores=0.54314774,
final_local_scores=0.5516867, pem=0.0, similarity=20.553144]
10: 740431 Mhow
    [pem_pos=0, final_scores=0.44543833,
final_local_scores=0.45131433, pem=0.0, similarity=17.557833]
11: 954785 Ratlam
    [pem_pos=0, final_scores=0.46374458,
final_local_scores=0.46722943, pem=-0.01594014,
similarity=18.119455]
12: 5706727 Ashta,_Madhya_Pradesh
    [pem_pos=0, final_scores=0.31659973,
final_local_scores=0.39372426, pem=-0.14310083,
similarity=16.617435]
13: 365670 Indore
    [pem_pos=0, final_scores=0.43150324,
final_local_scores=0.3904643, pem=-0.031899266,
similarity=15.91542]
14: 911235 Dhar
    [pem_pos=0, final_scores=0.51408625,
final_local_scores=0.5277174, pem=-0.210608,
similarity=20.983171]
15: 901117 Ujjain
    [pem_pos=0, final_scores=0.36304647,
final_local_scores=0.33074835, pem=0.0, similarity=13.959904]
16: 2001103 Jaora
    [pem_pos=0, final_scores=0.49458617,
final_local_scores=0.5601668, pem=0.0, similarity=20.806208]
```

```
17: 509811 Khandwa
    [pem_pos=0, final_scores=0.508715,
final_local_scores=0.5069734, pem=0.0, similarity=19.21881]
18: 2293572 Ashoknagar
    [pem_pos=0, final_scores=0.33988947,
final_local_scores=0.3406163, pem=0.0, similarity=14.254382]
19: 19189 Mumbai
    [pem_pos=0, final_scores=0.38104844,
final_local_scores=0.35429174, pem=-0.24257013,
similarity=15.981618]
```

Sample: b'603_Bangladesh'

[1 1 Bangladesh 1] Speaker says he received death threats . [2 2 DHAKA 2] 1996-08-27 The Speaker of [3 3 Bangladesh 3] 's parliament , Humayun Rasheed Choudhury , said he had received death threats from anonymous callers after opposition parties threatened to boycott proceedings chaired by him . He told the [4 4 Bengali 4] newspaper Banglabazar Patrika on Tuesday that such threats were possibly coming from " those who want to push the country into chaos and unrest . " The callers said his life could be cut short , the newspaper said . The speaker was not immediately available for comment . [5 5 Chowdhury 5] , a former foreign minister and veteran diplomat , was appointed speaker of the 330-member parliament on July 13 , a month after general elections returned the [6 6 Awami League 6] of Prime Minister Sheikh [7 7 Hasina 7] to power after 21 years . Choudhury also was president of the 41st session of the U.N. General Assembly in 1986-87 . Former prime minister [8 8 Begum Khaleda Zia 8] , now the opposition leader in parliament and head of the [9 9 Bangladesh Nationalist Party 9] ([10 10 BNP 10]) , said her followers might boycott assembly sessions chaired by the " partisan " speaker . " The ruling Awami league is making parliament ineffective and the speaker is contributing to that by not allowing the opposition MPs enough time to speak , " she told a rally in northern district of [11 11 Bogra 11] on Monday . [12 12 Hasina 12] , speaking to a group of engineers in [13 13 Dhaka 13] on Monday , accused the [14 14 BNP 14] of resorting to " terrorism " as part of its plan to create instability and chaos in the country . " This is not desireable ... and we will deal with such designs sternly , " the prime minister said .

```
1: 3454 Bangladesh
2: 56656 Dhaka
3: 3454 Bangladesh
4: 3454 Bangladesh
5: 12804134 Humayun_Rashid_Choudhury
6: 404861 Bangladesh_Awami_League
7: 623439 Sheikh_Hasina
8: 393763 Khaleda_Zia
9: 309760 Bangladesh_Nationalist_Party
10: 309760 Bangladesh_Nationalist_Party
```

```
11: 2247892 Bogra
12: 623439 Sheikh_Hasina
13: 56656 Dhaka
14: 309760 Bangladesh_Nationalist_Party
1: 3454 Bangladesh
    [pem_pos=0, final_scores=0.50696987,
final_local_scores=0.48779368, pem=-0.28971618,
similarity=20.221972]
2: 56656 Dhaka
    [pem_pos=0, final_scores=0.43552178,
final_local_scores=0.41209894, pem=-0.021151643,
similarity=16.502594]
3: 3454 Bangladesh
    [pem_pos=0, final_scores=0.5634804,
final_local_scores=0.5495228, pem=-0.28971618,
similarity=22.064093]
4: 3454 Bangladesh
    [pem_pos=6, final_scores=0.2836585,
final_local_scores=0.24385989, pem=-2.8055258,
similarity=26.623863]
5: 12804134 Humayun_Rashid_Choudhury
    [pem_pos=0, final_scores=0.36461496,
final_local_scores=0.44330996, pem=0.0, similarity=17.318966]
6: 404861 Bangladesh_Awami_League
    [pem_pos=0, final_scores=0.39085466,
final_local_scores=0.36033663, pem=0.0, similarity=14.842876]
7: 623439 Sheikh_Hasina
    [pem_pos=0, final_scores=0.4646774,
final_local_scores=0.44465983, pem=0.0, similarity=17.35925]
8: 393763 Khaleda_Zia
    [pem_pos=0, final_scores=0.5594388,
final_local_scores=0.54683596, pem=0.0, similarity=20.408388]
9: 309760 Bangladesh_Nationalist_Party
    [pem_pos=0, final_scores=0.60554075,
final_local_scores=0.5969011, pem=0.0, similarity=21.902433]
10: 309760 Bangladesh_Nationalist_Party
    [pem_pos=1, final_scores=0.4636855,
final_local_scores=0.44194585, pem=-1.374355,
similarity=24.752214]
11: 2247892 Bogra
    [pem_pos=0, final_scores=0.30665475,
final_local_scores=0.31233057, pem=0.0, similarity=13.41028]
12: 623439 Sheikh_Hasina
    [pem_pos=0, final_scores=0.534334,
final_local_scores=0.5207491, pem=0.0, similarity=19.629904]
13: 56656 Dhaka
    [pem_pos=0, final_scores=0.38513088,
final_local_scores=0.3570545, pem=-0.021151643,
similarity=14.859958]
14: 309760 Bangladesh_Nationalist_Party
    [pem_pos=1, final_scores=0.45682716,
final_local_scores=0.43445414, pem=-1.374355,
```

similarity=24.528648]

Sample: b'604_Bangladesh'

[1 1 Bangladesh 1] June M2 up 3.8 pct m / m , up 8.2 pct y / y . [2 2 DHAKA 2] 1996-08-27 [3 3 Bangladesh 3] 's M2 money supply rose 3.8 percent in June to 456.8 billion taka after a 0.27 percent rise to 439.9 billion in May , central bank officials said . The year-on-year rise was 8.2 percent to June , 1996 . [4 4 BANGLADESH 4] 'S MONEY SUPPLY JUNE MAY JUNE 1995 M2 money supply (bln taka) 456.8 439.9 422.1 M1 money supply (bln taka) 144.5 139.3 131.7
1: 3454 Bangladesh
2: 56656 Dhaka
3: 3454 Bangladesh
4: 3454 Bangladesh
1: 3454 Bangladesh
[pem_pos=0, final_scores=0.60424286, final_local_scores=0.5376455, pem=-0.28971618, similarity=21.70965]
2: 56656 Dhaka
[pem_pos=0, final_scores=0.48348665, final_local_scores=0.4513082, pem=-0.021151643, similarity=17.672676]
3: 3454 Bangladesh
[pem_pos=0, final_scores=0.5816311, final_local_scores=0.51294565, pem=-0.28971618, similarity=20.972557]
4: 3454 Bangladesh
[pem_pos=0, final_scores=0.50188607, final_local_scores=0.42583627, pem=-0.28971618, similarity=18.37304]

Sample: b'605_HELIBOR'

[1 1 HELIBOR 1] INTEREST RATES LARGELY UNCHANGED . [2 2 HELS INKI 2] 1996-08-27 [3 3 Helibor 3] market interest rates were largely unchanged at the [4 4 Bank of Finland 4] 's daily fixing on Tuesday . The key three-month rate was steady at 3.40 percent . August 27 fix August 26 fix 1-mth [5 5 Helibor 5] 3.27 pct 3.29 pct 2-mth [6 6 Helibor 6] 3.34 pct 3.34 pct 3-mth [7 7 Helibor 7] 3.40 pct 3.40 pct 6-mth [8 8 Helibor 8] 3.56 pct 3.55 pct 9-mth [9 9 Helibor 9] 3.73 pct 3.70 pct 12-mth [10 10 Helibor 10] 3.89 pct 3.87 pct -- [11 11 Helsinki 11] newsroom +358 - 0 - 680 50 248

1: 21112335 Helibor
2: 13696 Helsinki
3: 21112335 Helibor
4: 850543 Bank_of_Finland
5: 21112335 Helibor
6: 21112335 Helibor
7: 21112335 Helibor
8: 21112335 Helibor
9: 21112335 Helibor

```
10: 21112335 Helibor
11: 13696 Helsinki
1: 21112335 Helibor
    [pem_pos=0, final_scores=0.40356302,
final_local_scores=0.3299495, pem=0.0, similarity=13.936064]
2: 13696 Helsinki
    [pem_pos=0, final_scores=0.48745877,
final_local_scores=0.62930757, pem=-0.015504186,
similarity=22.953821]
3: 21112335 Helibor
    [pem_pos=0, final_scores=0.5485473,
final_local_scores=0.4883228, pem=0.0, similarity=18.66224]
4: 850543 Bank_of_Finland
    [pem_pos=0, final_scores=0.58972484,
final_local_scores=0.7082113, pem=0.0, similarity=25.224152]
5: 21112335 Helibor
    [pem_pos=0, final_scores=0.6782983,
final_local_scores=0.6300561, pem=0.0, similarity=22.891844]
6: 21112335 Helibor
    [pem_pos=0, final_scores=0.7421656,
final_local_scores=0.6998214, pem=0.0, similarity=24.97378]
7: 21112335 Helibor
    [pem_pos=0, final_scores=0.7293164,
final_local_scores=0.68578553, pem=0.0, similarity=24.55492]
8: 21112335 Helibor
    [pem_pos=0, final_scores=0.73557585,
final_local_scores=0.6926231, pem=0.0, similarity=24.758966]
9: 21112335 Helibor
    [pem_pos=0, final_scores=0.7263494,
final_local_scores=0.68254465, pem=0.0, similarity=24.458206]
10: 21112335 Helibor
    [pem_pos=0, final_scores=0.71324486,
final_local_scores=0.6682298, pem=0.0, similarity=24.031023]
11: 13696 Helsinki
    [pem_pos=0, final_scores=0.48395866,
final_local_scores=0.6254842, pem=-0.015504186,
similarity=22.839725]
```

Sample: b'606_Barrick'
[1 1 Barrick 1] gets 93 pct of Arequipa . [2 2 TORONTO 2]
1996-08-27 Barrick Gold Corp said on Tuesday its takeover
offer for Arequipa Resources Ltd was successful , with 93
percent of the 36.3 million shares not already owned tendered
under the bid , which expired overnight . " We are pleased
that Arequipa shareholders ahave chosen so overwhelmingly to
accept this offer . We now have the opportunity to realize
the potential of Arequipa 's excellent assets , " [3 3 Barrick
3] chairman and chief executive [4 4 Peter Munk 4] said in a
statement . The C\$ 30-a-share deal means [5 5 Barrick 5] will
own Arequipa 's attractive Pierina gold deposit in [6 6 Peru 6
] . [7 7 Barrick 7] said details involving the allocation
between [8 8 Barrick 8] shares and cash will be available

shortly . [9 9 Barrick 9] 's offer of C\$ 30 a share or part cash , part share offer was [10 10 Barrick 10] 's second attempt to swallow the small [11 11 Vancouver-based 11] gold prospector . [12 12 Toronto-based 12] [13 13 Barrick 13] , the world 's third largest gold producer , sweetened its July 11 bid to C\$ 30 a share from C\$ 27 on August 16 after a fresh batch of drill results from the Pierina deposit . Experts have speculated the deposit has potential reserves of up to 12 million ounces . More drilling results are expected soon . The [14 14 Barrick 14] bid took observers by surprise , since Arequipa 's exploration was still in its early stages . Arequipa shareholders had the option to choose C\$ 30 cash or 0.79 [15 15 Barrick 15] shares plus 50 cents for each Arequipa share . Shares were to be pro-rated if more than 14.4 million were requested . -- Reuters Toronto Bureau 416 941-8100

```
1: 867647 Barrick_Gold
2: 64646 Toronto
3: 867647 Barrick_Gold
4: 979015 Peter_Munk
5: 867647 Barrick_Gold
6: 170691 Peru
7: 867647 Barrick_Gold
8: 867647 Barrick_Gold
9: 867647 Barrick_Gold
10: 867647 Barrick_Gold
11: 32706 Vancouver
12: 64646 Toronto
13: 867647 Barrick_Gold
14: 867647 Barrick_Gold
15: 867647 Barrick_Gold
1: 867647 Barrick_Gold
    [pem_pos=0, final_scores=0.8010177,
final_local_scores=0.7640441, pem=0.0, similarity=26.890312]
2: 64646 Toronto
    [pem_pos=0, final_scores=0.50570756,
final_local_scores=0.5726269, pem=-0.12291032,
similarity=21.846447]
3: 867647 Barrick_Gold
    [pem_pos=0, final_scores=0.8363009,
final_local_scores=0.80258566, pem=0.0, similarity=28.04047]
4: 979015 Peter_Munk
    [pem_pos=0, final_scores=0.32602912,
final_local_scores=0.40940928, pem=0.0, similarity=16.307302]
5: 867647 Barrick_Gold
    [pem_pos=0, final_scores=0.76218456,
final_local_scores=0.72162485, pem=0.0, similarity=25.624437]
6: 170691 Peru
    [pem_pos=0, final_scores=0.52197075,
final_local_scores=0.6182043, pem=-0.23383401,
similarity=23.809788]
7: 867647 Barrick_Gold
```

```
[pem_pos=0, final_scores=0.6795749,
final_local_scores=0.6313863, pem=0.0, similarity=22.93154]
8: 867647 Barrick_Gold
    [pem_pos=0, final_scores=0.93322664,
final_local_scores=0.90846235, pem=0.0, similarity=31.200043]
9: 867647 Barrick_Gold
    [pem_pos=0, final_scores=0.76798743,
final_local_scores=0.72796357, pem=0.0, similarity=25.813599]
10: 867647 Barrick_Gold
    [pem_pos=0, final_scores=0.75062627,
final_local_scores=0.70899916, pem=0.0, similarity=25.247662]
11: 32706 Vancouver
    [pem_pos=0, final_scores=0.57491654,
final_local_scores=0.6490082, pem=0.0, similarity=23.457413]
12: 64646 Toronto
    [pem_pos=0, final_scores=0.57053643,
final_local_scores=0.6434425, pem=0.0, similarity=23.29132]
13: 867647 Barrick_Gold
    [pem_pos=0, final_scores=0.9013414,
final_local_scores=0.8736325, pem=0.0, similarity=30.16065]
14: 867647 Barrick_Gold
    [pem_pos=0, final_scores=0.7609591,
final_local_scores=0.72028613, pem=0.0, similarity=25.584488]
15: 867647 Barrick_Gold
    [pem_pos=0, final_scores=0.8609248,
final_local_scores=0.82948345, pem=0.0, similarity=28.843153]
```

Sample: b'607_Penn'

Penn Treaty terminates acquisition pact . [1 1 ALLENTOWN
1] , Pa . 1996-08-27 Penn Treaty American Corp said Tuesday it terminated a previously announced non-binding letter of intent to purchase Merrion Insurance Company Inc , a [2 2 New York 2] licensed company . In announcing its decision , Penn Treaty said it " will continue to actively pursue entering into the [3 3 New York 3] long-term care market through licensing or by acquisition . " It explained the " addition of a [4 4 New York 4] license will enable Penn Treaty American Corp to conduct business in all 50 states , following the company 's acquisition of Health Insurance of Vermont , a Vermont domiciled insurer , scheduled to close on August 30 , 1996 . " -- New York Newsdesk 212-859-1610 .

```
1: 132982 Allentown,_Pennsylvania
2: 645042 New_York_City
3: 645042 New_York_City
4: 645042 New_York_City
1: 132982 Allentown,_Pennsylvania
    [pem_pos=0, final_scores=0.60251355,
final_local_scores=0.61859566, pem=-0.32889923,
similarity=24.338448]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.5029078,
final_local_scores=0.4310825, pem=-1.7906687,
```

```
similarity=26.692009]
3: 645042 New_York_City
    [pem_pos=1, final_scores=0.44808406,
final_local_scores=0.37119585, pem=-1.7906687,
similarity=24.904873]
4: 645042 New_York_City
    [pem_pos=1, final_scores=0.50059706,
final_local_scores=0.4285583, pem=-1.7906687,
similarity=26.616682]

Sample: b'608_VNU'
VNU details first-half operating profits . [1 1 HAARLEM 1] ,
[2 2 Netherlands 2] 1996-08-27 Publisher VNU gave the
following breakdown of its first-half results : H1 1996 H1
1995 Sales Op profit Sales Op profit Consumer magazines 618
90 568 80 Newspapers 363 49 295 46 Commercial TV 127 6 loss
174 33 profit Business info [3 3 Europe 3] 231 24 178 18
Business info [4 4 USA 4] 382 61 362 41 Education 42 6 36 3
Miscellaneous charges --- 16 --- 30 NOTES - Sales and
operating profit are given in millions of guilders .
Commercial TV includes pro rata share of sales and operating
profits in [5 5 Dutch 5] group HMG and [6 6 Belgium 6] 's [7 7
VTM 7] . -- [8 8 Amsterdam 8] newsroom +31 20 504 5000 , Fax
+31 20 504 5040
1: 63278 Haarlem
2: 21148 Netherlands
3: 9239 Europe
4: 3434750 United_States
5: 21148 Netherlands
6: 3343 Belgium
7: 1165422 Vtm
8: 844 Amsterdam
1: 63278 Haarlem
    [pem_pos=0, final_scores=0.46190983,
final_local_scores=0.48080468, pem=0.0, similarity=18.437883]
2: 21148 Netherlands
    [pem_pos=0, final_scores=0.6137714,
final_local_scores=0.5627142, pem=-0.24940795,
similarity=22.238546]
3: 9239 Europe
    [pem_pos=0, final_scores=0.443941,
final_local_scores=0.45646906, pem=-0.086338006,
similarity=18.181181]
4: 3434750 United_States
    [pem_pos=0, final_scores=0.4423719,
final_local_scores=0.4528185, pem=-0.13038418,
similarity=18.311771]
5: 21148 Netherlands
    [pem_pos=0, final_scores=0.6766081,
final_local_scores=0.6313538, pem=-0.7943452,
similarity=27.250341]
6: 3343 Belgium
```

```
[pem_pos=0, final_scores=0.6392154,
final_local_scores=0.6048909, pem=-0.24180195,
similarity=23.45582]
7: 1165422 Vtm
    [pem_pos=0, final_scores=0.30433434,
final_local_scores=0.42086416, pem=0.0, similarity=16.64914]
8: 844 Amsterdam
    [pem_pos=0, final_scores=0.6159598,
final_local_scores=0.59043986, pem=-0.0642803,
similarity=22.059183]
```

Sample: b'609_AOL'

AOL Europe forms online advertising agency . [1 1 HANOVER 1] , [2 2 Germany 2] 1996-08-27 The joint venture between [3 3 America Online 3] ([4 4 AOL 4]) and [5 5 Bertelsmann AG 5] has formed a new company to sell advertising space on [6 6 AOL 6] in [7 7 Europe 7] , a [8 8 Bertelsmann 8] official said on Tuesday . The new company is called AdOn GmbH and is located in [9 9 Hamburg 9] , Bernd Schiphorst , president and chief operating officer of [10 10 Bertelsmann 10] 's [11 11 New Media 11] business division said on the sidelines of an [12 12 AOL 12] news conference at the [13 13 CeBIT 13] Home consumer electronics fair in Hanover . Jan Buettner , managing director of AOL Germany , said AdOn was in the formation phase and would " get off the ground next year , bringing in advertising from around [14 14 Europe 14] . " He would give no forecasts on advertising revenue . [15 15 AOL 15] is the leading global commercial online service with some six million subscribers worldwide . It has around 200,000 subscribers in [16 16 Europe 16] , with two thirds of that number in [17 17 Germany 17] alone . In [18 18 Europe 18] , the service is available in [19 19 Germany 19] , [20 20 France 20] and [21 21 Britain 21] . It will be available in [22 22 Austria 22] and [23 23 Switzerland 23] later this year and in Scandinavia and the [24 24 Benelux 24] countries next year . -- William Boston , [25 25 CeBIT 25] newsroom , 0172 6736510

```
1: 14197 Hanover
2: 11867 Germany
3: 1397 AOL
4: 1397 AOL
5: 140740 Bertelsmann
6: 1397 AOL
7: 9239 Europe
8: 140740 Bertelsmann
9: 13467 Hamburg
10: 140740 Bertelsmann
11: 344859 New_media
12: 1397 AOL
13: 25079983 CeBIT
14: 9239 Europe
15: 1397 AOL
16: 9239 Europe
```

17: 11867 Germany
18: 9239 Europe
19: 11867 Germany
20: 5843419 France
21: 31717 United_Kingdom
22: 26964606 Austria
23: 26748 Switzerland
24: 4170 Benelux
25: 25079983 CeBIT
1: 14197 Hanover
 [pem_pos=0, final_scores=0.48775798,
final_local_scores=0.50340754, pem=-0.7071154,
similarity=22.957802]
2: 11867 Germany
 [pem_pos=0, final_scores=0.8414107,
final_local_scores=0.8459169, pem=-0.15529114,
similarity=30.17806]
3: 1397 AOL
 [pem_pos=0, final_scores=0.5256167,
final_local_scores=0.55292875, pem=0.0, similarity=20.59021]
4: 1397 AOL
 [pem_pos=0, final_scores=0.44098097,
final_local_scores=0.46047705, pem=0.0, similarity=17.831266]
5: 140740 Bertelsmann
 [pem_pos=0, final_scores=0.39849216,
final_local_scores=0.4200731, pem=0.0, similarity=16.625532]
6: 1397 AOL
 [pem_pos=0, final_scores=0.57084537,
final_local_scores=0.60233414, pem=0.0, similarity=22.064566]
7: 9239 Europe
 [pem_pos=0, final_scores=0.6419008,
final_local_scores=0.6468981, pem=-0.086338006,
similarity=23.863962]
8: 140740 Bertelsmann
 [pem_pos=0, final_scores=0.25841433,
final_local_scores=0.26705936, pem=0.0, similarity=12.059297]
9: 13467 Hamburg
 [pem_pos=0, final_scores=0.40820032,
final_local_scores=0.40540606, pem=-0.27202022,
similarity=17.667126]
10: 140740 Bertelsmann
 [pem_pos=0, final_scores=0.4035356,
final_local_scores=0.4255823, pem=0.0, similarity=16.789938]
11: 344859 New_media
 [pem_pos=0, final_scores=0.44190735,
final_local_scores=0.48440176, pem=0.0, similarity=18.545227]
12: 1397 AOL
 [pem_pos=0, final_scores=0.5882592,
final_local_scores=0.6213561, pem=0.0, similarity=22.63222]
13: 25079983 CeBIT
 [pem_pos=0, final_scores=0.54398686,
final_local_scores=0.6232974, pem=0.0, similarity=22.690151]

```
14: 9239 Europe
    [pem_pos=0, final_scores=0.6461569,
final_local_scores=0.65154725, pem=-0.086338006,
similarity=24.002703]
15: 1397 AOL
    [pem_pos=0, final_scores=0.4471566,
final_local_scores=0.467223, pem=0.0, similarity=18.03258]
16: 9239 Europe
    [pem_pos=0, final_scores=0.60143185,
final_local_scores=0.60269195, pem=-0.086338006,
similarity=22.544764]
17: 11867 Germany
    [pem_pos=0, final_scores=0.70863026,
final_local_scores=0.7008744, pem=-0.15529114,
similarity=25.8497]
18: 9239 Europe
    [pem_pos=0, final_scores=0.59114176,
final_local_scores=0.5914516, pem=-0.086338006,
similarity=22.209328]
19: 11867 Germany
    [pem_pos=0, final_scores=0.6231515,
final_local_scores=0.6075018, pem=-0.15529114,
similarity=23.063276]
20: 5843419 France
    [pem_pos=0, final_scores=0.6708722,
final_local_scores=0.69300085, pem=-0.16240834,
similarity=25.653442]
21: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.6051884,
final_local_scores=0.6030471, pem=-0.61114657,
similarity=25.40935]
22: 26964606 Austria
    [pem_pos=0, final_scores=0.46498907,
final_local_scores=0.44465965, pem=-0.25262418,
similarity=18.733053]
23: 26748 Switzerland
    [pem_pos=0, final_scores=0.5156003,
final_local_scores=0.51021665, pem=-0.18598008,
similarity=20.326984]
24: 4170 Benelux
    [pem_pos=0, final_scores=0.43553454,
final_local_scores=0.4407289, pem=0.0, similarity=17.241943]
25: 25079983 CeBIT
    [pem_pos=0, final_scores=0.3944599,
final_local_scores=0.4599619, pem=0.0, similarity=17.815893]
```

Sample: b'610_All'

All passengers freed from [1 1 Sudanese 1] hijack plane . [2 2 STANSTED 2] , [3 3 England 3] 1996-08-27 All passengers held hostage aboard a hijacked [4 4 Sudanese 4] Airways plane diverted to [5 5 London 5] 's [6 6 Stansted 6] airport carrying 199 passengers and crew have been freed , an airport

spokeswoman said on Tuesday . Eyewitnesses said six crew members had also been allowed to leave the aircraft . Airport spokeswoman Rona Young confirmed that all the passengers had left the aircraft . Police said a number of crew members had left the aircraft and said details would be given at a news conference expected to be held in the next few minutes by the local police chief . The passengers were released in batches during the course of the morning after the Airbus A310 landed at [7 7 Stansted 7] , having been diverted from [8 8 Cyprus 8] . The aircraft was hijacked on a flight from [9 9 Khartoum 9] to [10 10 Amman 10] by six or seven men , who police say may be [11 11 Iraqis 11] . -- London Newsroom +
00--44-171-542-7947

1: 27421 Sudan
2: 65456 London_Stansted_Airport
3: 9316 England
4: 27421 Sudan
5: 17867 London
6: 65456 London_Stansted_Airport
7: 65456 London_Stansted_Airport
8: 5593 Cyprus
9: 17320 Khartoum
10: 45007 Amman
11: 7515928 Iraq
1: 27421 Sudan
[pem_pos=0, final_scores=0.5099083,
final_local_scores=0.49853814, pem=-0.19893567,
similarity=20.04893]
2: 65456 London_Stansted_Airport
[pem_pos=0, final_scores=0.34470063,
final_local_scores=0.35497153, pem=-0.2670293,
similarity=16.134918]
3: 9316 England
[pem_pos=0, final_scores=0.61402684,
final_local_scores=0.6392899, pem=-0.19053856,
similarity=24.203577]
4: 27421 Sudan
[pem_pos=0, final_scores=0.45958668,
final_local_scores=0.44356936, pem=-0.19893567,
similarity=18.40855]
5: 17867 London
[pem_pos=0, final_scores=0.58920306,
final_local_scores=0.5948686, pem=-0.09086737,
similarity=22.335928]
6: 65456 London_Stansted_Airport
[pem_pos=0, final_scores=0.43210417,
final_local_scores=0.45044667, pem=-0.2670293,
similarity=18.984087]
7: 65456 London_Stansted_Airport
[pem_pos=0, final_scores=0.38612908,
final_local_scores=0.40022582, pem=-0.2670293,
similarity=17.485395]

```
8: 5593 Cyprus
    [pem_pos=0, final_scores=0.553131,
final_local_scores=0.5610976, pem=-0.3752463,
similarity=22.874632]
9: 17320 Khartoum
    [pem_pos=0, final_scores=0.42001158,
final_local_scores=0.39416206, pem=-0.03966028,
similarity=16.067974]
10: 45007 Amman
    [pem_pos=0, final_scores=0.41370124,
final_local_scores=0.39817572, pem=0.0, similarity=15.972071]
11: 7515928 Iraq
    [pem_pos=1, final_scores=0.45844704,
final_local_scores=0.44959742, pem=-0.8129937,
similarity=21.927784]
```

Sample: b'611_Swiss'

```
[1 1 Swiss Bank Corp 1] sets warrants on DTB-Bund-Future . [2
2 LONDON 2] 1996-08-27 [3 3 Swiss Bank Corp 3] says it has
issued 60 million American-style call and put warrants , in
six equal tranches , on the DTB-Bund-Future March 1997 .
EXERCISE PERIOD 02.SEP.96-06.MAR.97 PAYDATE 30.AUG.96 LISTING
FFT DDF MIN EXER LOT 100 SPOT REFERENCE 95.35 PCT WARRANTS
STRIKE ISS PRICE PREMIUM GEARING CALL A 96.00 PCT 1.16 [4 4 DE
M 4] 1.90 PCT 82.20 X CALL B 97.00 PCT 0.75 [5 5 DEM 5] 2.50
PCT 127.10 X CALL C 98.00 PCT 0.47 [6 6 DEM 6] 3.30 PCT
202.90 X PUT D 94.00 PCT 0.94 [7 7 DEM 7] 2.40 PCT 101.40 X
PUT E 95.0 PCT 1.33 [8 8 DEM 8] 1.80 PCT 71.70 X PUT F 96.0
PCT 1.84 [9 9 DEM 9] 1.20 PCT 51.80 X -- Reuter London
Newsroom +44 171 542 7658
1: 520367 Swiss_Bank_Corporation
2: 17867 London
3: 520367 Swiss_Bank_Corporation
4: 88160 Deutsche_Mark
5: 88160 Deutsche_Mark
6: 88160 Deutsche_Mark
7: 88160 Deutsche_Mark
8: 88160 Deutsche_Mark
9: 88160 Deutsche_Mark
1: 520367 Swiss_Bank_Corporation
    [pem_pos=0, final_scores=0.37367976,
final_local_scores=0.4466936, pem=0.0, similarity=17.41994]
2: 17867 London
    [pem_pos=0, final_scores=0.51011795,
final_local_scores=0.6308503, pem=-0.09086737,
similarity=23.409695]
3: 520367 Swiss_Bank_Corporation
    [pem_pos=0, final_scores=0.3964414,
final_local_scores=0.47155726, pem=0.0, similarity=18.161922]
4: 88160 Deutsche_Mark
    [pem_pos=2, final_scores=0.8792862,
final_local_scores=0.8542189, pem=-1.9186858,
```

```
similarity=40.01542]
5: 88160 Deutsche_Mark
    [pem_pos=2, final_scores=0.844353,
final_local_scores=0.81605965, pem=-1.9186858,
similarity=38.87667]
6: 88160 Deutsche_Mark
    [pem_pos=2, final_scores=0.86800534,
final_local_scores=0.84189624, pem=-1.9186858,
similarity=39.647686]
7: 88160 Deutsche_Mark
    [pem_pos=2, final_scores=0.88465005,
final_local_scores=0.86007804, pem=-1.9186858,
similarity=40.19027]
8: 88160 Deutsche_Mark
    [pem_pos=2, final_scores=0.88733584,
final_local_scores=0.8630118, pem=-1.9186858,
similarity=40.277817]
9: 88160 Deutsche_Mark
    [pem_pos=2, final_scores=0.88287526,
final_local_scores=0.85813934, pem=-1.9186858,
similarity=40.132416]
```

Sample: b'612_DBRS'

```
[1 1 DBRS 1] confirms Power Corp , Power Financial ratings .
[2 2 TORONTO 2] 1996-08-27 [3 3 Dominion Bond Rating Service 3]
] said on Tuesday it confirmed the ratings on Power Corp of
Canada 's senior debt and preferred shares at A ( high ) and
Pfd-2 , respectively , with stable trends . [4 4 DBRS 4] said
it also confirmed Power Financial Corp 's senior debentures ,
cumulative preferred shares and non-cumulative first
preferred shares , series B , at AA ( low ) , Pfd-1 and Pfd-1
( low ) , all with stable trends .
1: 4526656 Dominion_Bond_Rating_Service
2: 64646 Toronto
3: 4526656 Dominion_Bond_Rating_Service
4: 4526656 Dominion_Bond_Rating_Service
1: 4526656 Dominion_Bond_Rating_Service
    [pem_pos=0, final_scores=0.5417047,
final_local_scores=0.48968756, pem=0.0, similarity=18.702967]
2: 64646 Toronto
    [pem_pos=0, final_scores=0.42888314,
final_local_scores=0.5778289, pem=-0.12291032,
similarity=22.001686]
3: 4526656 Dominion_Bond_Rating_Service
    [pem_pos=0, final_scores=0.51978165,
final_local_scores=0.46573997, pem=0.0, similarity=17.988323]
4: 4526656 Dominion_Bond_Rating_Service
    [pem_pos=0, final_scores=0.5702111,
final_local_scores=0.5208265, pem=0.0, similarity=19.632215]
```

Sample: b'613_Turkey'

```
Turkey 's [1 1 Kurd 1] rebels kill two , take three
```

hostage . [2 2 ANKARA 2] 1996-08-27 [3 3 Kurdish 3]
guerrillas killed two people and took three hostage after
stopping two intercity buses at a roadblock in eastern
Turkey , security officials said on Tuesday . They told
reporters a group of [4 4 Kurdistan Workers Party 4] ([5 5 PK
K 5]) guerrillas stopped the buses at a roadblock on the
road linking the eastern provinces of [6 6 Erzincan 6] and [7
7 Sivas 7] on Monday night and forced the passengers to get
out . The rebels killed one of the drivers and a passenger
after checking the identities of the passengers , they said .
The officials said the rebels set ablaze two buses and
released all but three passengers . More than 20,000 people
have been killed in the 12-year-old conflict between Turkish
troops and [8 8 PKK 8] guerrillas fighting for autonomy or
independence from Turkey .

1: 80777 Kurdistan
2: 802 Ankara
3: 80777 Kurdistan
4: 69680 Kurdistan_Workers'_Party
5: 69680 Kurdistan_Workers'_Party
6: 1210336 Erzincan
7: 1376207 Sivas
8: 69680 Kurdistan_Workers'_Party
1: 80777 Kurdistan
[pem_pos=3, final_scores=0.32405442,
final_local_scores=0.29052544, pem=-2.608398,
similarity=26.944445]
2: 802 Ankara
[pem_pos=0, final_scores=0.40280265,
final_local_scores=0.40098572, pem=0.0, similarity=16.055927]
3: 80777 Kurdistan
[pem_pos=9, final_scores=0.38368118,
final_local_scores=0.35565865, pem=-2.8001814,
similarity=29.931097]
4: 69680 Kurdistan_Workers'_Party
[pem_pos=0, final_scores=0.6011545,
final_local_scores=0.5839382, pem=0.0, similarity=21.515593]
5: 69680 Kurdistan_Workers'_Party
[pem_pos=0, final_scores=0.5808508,
final_local_scores=0.5617594, pem=0.0, similarity=20.853735]
6: 1210336 Erzincan
[pem_pos=0, final_scores=0.3291958,
final_local_scores=0.31846625, pem=-0.25802472,
similarity=14.996559]
7: 1376207 Sivas
[pem_pos=0, final_scores=0.31882054,
final_local_scores=0.30621698, pem=0.0, similarity=13.2278385]
8: 69680 Kurdistan_Workers'_Party
[pem_pos=0, final_scores=0.5426159,
final_local_scores=0.5199936, pem=0.0, similarity=19.607359]

Sample: b'614_Egypt'

[1 1 Egypt 1] confiscates paper for " mad rulers " article . [2 2 CAIRO 2] 1996-08-27 [3 3 Egypt 3] has banned and confiscated 10,000 copies of the Cyprus-based [4 4 Arabic 4] monthly newspaper al-Tadamun because of an editorial suggesting mental health tests for [5 5 Arab 5] leaders , the editor-in-chief said on Tuesday . Mohamed Abu Liwaya , said Information Ministry censors had told him to send all the copies of the August edition back to [6 6 Cyprus 6] at his own expense . He told [7 7 Reuters 7] the reason was his own front-page editorial , entitled " A Chronic Mental Illness " in which he attacks compliant [8 8 Arab 8] leaders for serving [9 9 U.S. 9] and [10 10 Israeli 10] interests . " The [11 11 Arabs 11] demand that our [12 12 Arab 12] leaders undergo a compulsory examination by a team of psychiatrists to see how sound their mental capacities are , " the editorial said . " Because our leaders have started to behave with extreme hostility towards the interests of their peoples to court the goodwill of the [13 13 Americans 13] and the [14 14 Zionists 14] , " he added . The censorship office denied they had confiscated the newspapers but declined to say when they could go on sale .

1: 8087628 Egypt

2: 6293 Cairo

3: 8087628 Egypt

4: 2185 Arab_people

5: 2185 Arab_people

6: 5593 Cyprus

7: 18998750 Reuters

8: 2185 Arab_people

9: 3434750 United_States

10: 9282173 Israel

11: 2185 Arab_people

12: 2185 Arab_people

13: 3434750 United_States

14: 34484 Zionism

1: 8087628 Egypt

[pem_pos=0, final_scores=0.5116613,
final_local_scores=0.4599808, pem=-0.27480945,
similarity=19.310913]

2: 6293 Cairo

[pem_pos=0, final_scores=0.57159513,
final_local_scores=0.5559701, pem=-0.15211283,
similarity=21.508183]

3: 8087628 Egypt

[pem_pos=0, final_scores=0.56659085,
final_local_scores=0.51998293, pem=-0.27480945,
similarity=21.101498]

4: 2185 Arab_people

[pem_pos=7, final_scores=0.12844719,
final_local_scores=0.044619203, pem=-3.4843605,
similarity=24.369734]

5: 2185 Arab_people

```
[pem_pos=1, final_scores=0.40990567,
final_local_scores=0.35056415, pem=-2.0453336,
similarity=25.67409]
6: 5593 Cyprus
    [pem_pos=0, final_scores=0.6260795,
final_local_scores=0.63171214, pem=-0.3752463,
similarity=24.98191]
7: 18998750 Reuters
    [pem_pos=0, final_scores=0.38586807,
final_local_scores=0.41825223, pem=0.0, similarity=16.571194]
8: 2185 Arab_people
    [pem_pos=1, final_scores=0.34788465,
final_local_scores=0.2828156, pem=-2.0453336,
similarity=23.652338]
9: 3434750 United_States
    [pem_pos=0, final_scores=0.5662214,
final_local_scores=0.5825756, pem=-0.14075686,
similarity=22.240389]
10: 9282173 Israel
    [pem_pos=0, final_scores=0.53022206,
final_local_scores=0.5062143, pem=-0.34790626,
similarity=21.088125]
11: 2185 Arab_people
    [pem_pos=1, final_scores=0.37265402,
final_local_scores=0.30987236, pem=-1.8745545,
similarity=23.53104]
12: 2185 Arab_people
    [pem_pos=1, final_scores=0.3420117,
final_local_scores=0.2764003, pem=-2.0453336,
similarity=23.460892]
13: 3434750 United_States
    [pem_pos=0, final_scores=0.45234722,
final_local_scores=0.45818537, pem=-0.5581965,
similarity=20.798437]
14: 34484 Zionism
    [pem_pos=0, final_scores=0.32274848,
final_local_scores=0.33366033, pem=0.0, similarity=14.0468025]
```

Sample: b'615_IPO'

[1 1 IPO 1] FILING - Transkaryotic Therapies Inc . [2 2 WASH
INGTON 2] 1996-08-27 Company Name Transkaryotic Therapies Inc
[3 3 Nasdaq 3] Stock symbol TKTX Estimated price range \$ 13 -
\$ 15 / shr Total shares to be offered 2.5 million Shrs
offered by company 2.5 million Shrs outstanding after ipo
16,668,560 Lead Underwriter Morgan Stanley and Co Inc
Underwriters over-allotment 375,000 shrs Shares to be
purchased by Hoechst Marion Roussel Inc 357,143 Business :
developed two proprietary technology platforms , gene
activation and gene therapy . Use of Proceeds : Research ,
preclinical and clinical product development , and general
corporate purposes . Financial Data in 000s : 1995 1994 -
Revenue \$ 15,400 \$ 10,000 - Net Income (loss) \$ 2,074 (\$

```
3,422 )
1: 4062633 Intellectual_Property_Owners_Association
2: 108956 Washington,_D.C.
3: 21559 NASDAQ
1: 4062633 Intellectual_Property_Owners_Association
    [pem_pos=4, final_scores=0.33528823,
final_local_scores=0.4978506, pem=-1.9454234,
similarity=29.526081]
2: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.44113535,
final_local_scores=0.5660468, pem=-1.7496481,
similarity=30.496536]
3: 21559 NASDAQ
    [pem_pos=0, final_scores=0.5811775,
final_local_scores=0.7095187, pem=0.0, similarity=25.263165]
```

Sample: b'616_AMTRAK'

[1 1 AMTRAK 1] train hits truck , derails in Vermont . [2 2 M
ONTPELIER 2] , Vt . 1996-08-27 An [3 3 Amtrak 3] train struck
a logging truck early on Tuesday and derailed , Vermont state
police said . A state police spokeswoman said there were
reports of minor injuries as a result of the derailment near [
4 4 Roxbury 4] , a small town on the edge of the Northfield
Mountains some 15 miles southwest of [5 5 Montpelier 5] , the
state capital . Further details were not immediately
available .

```
1: 51928 Amtrak
2: 41992 Montpelier,_Vermont
3: 51928 Amtrak
4: 137418 Roxbury,_Vermont
5: 41992 Montpelier,_Vermont
1: 51928 Amtrak
    [pem_pos=0, final_scores=0.42148858,
final_local_scores=0.4269781, pem=0.0, similarity=16.83159]
2: 41992 Montpelier,_Vermont
    [pem_pos=0, final_scores=0.4235335,
final_local_scores=0.4117785, pem=-0.89688367,
similarity=21.255398]
3: 51928 Amtrak
    [pem_pos=0, final_scores=0.3451404,
final_local_scores=0.3435793, pem=0.0, similarity=14.342804]
4: 137418 Roxbury,_Vermont
    [pem_pos=4, final_scores=0.29635984,
final_local_scores=0.37005192, pem=-2.104593,
similarity=26.577904]
5: 41992 Montpelier,_Vermont
    [pem_pos=0, final_scores=0.42100662,
final_local_scores=0.40901828, pem=-0.89688367,
similarity=21.173027]
```

Sample: b'617_Wife'

Wife of gun victim [1 1 Brady 1] praises [2 2 Clinton 2] . [3

3 CHICAGO 3] 1996-08-26 [4 4 Sarah Brady 4] , whose [5 5 Repu blican 5] husband was severely disabled in an assassination attempt on President [6 6 Ronald Reagan 6] , took centre stage at the Democratic National convention on Monday night to praise President [7 7 Bill Clinton 7] 's gun control efforts . With her husband 1 James sitting in a wheelchair to the side of the podium , Mrs. [8 8 Brady 8] called the handgun control bill that a [9 9 Democratic 9] [10 10 Congress 10] passed and [11 11 Clinton 11] signed in 1994 a major step in controlling firearm violence in the [12 12 United States 12] . But she said more had to be done . The Bradys walked on to the stage , he on her arm and with the aid of a cane , to a rousing reception from the convention . Their teenaged son sat in a VIP box with first lady [13 13 Hillary Rodham Clinton 13] and watched as his father returned to his wheelchair . " Jim , we must have made a wrong turn . This is n't [14 14 San Diego 14] (site of the [15 15 Republican 15] convention) , " Mrs. Brady joked to her husband , who was serving as [16 16 Reagan 16] 's press secretary when he was shot . " Sarah , I told you this is the [17 17 Democratic 17] convention , " he responded to his wife , who before the shooting had worked for two [18 18 Republican 18] congressmen and the [19 19 Repu blican 19] national party . " Since the Brady Law went into effect on February 28 , 1994 (it) has stopped more than 100,000 convicted felons and other prohibited purchasers from buying a handgun . Today , and every day , the Brady Law is stopping an estimated 85 felons from buying a handgun , " Mrs. [20 20 Brady 20] said . She added , " But we need to do more . We should , as President [21 21 Clinton 21] proposed today , stop people convicted of domestic violence from buying a handgun . Jim and I join with you tonight in saluting the great job that President [22 22 Clinton 22] has done in fighting crime and gun violence . " " He 's a hunter and a sportsman , but he understands the difference between a [23 23 Remington 23] rifle and an [24 24 AK47 24] . And he knows that you do n't go hunting with an [25 25 Uzi 25] . Mr. President you deserve our thanks . " [26 26 Jim Brady 26] then gave a big thumbs to the audience . [27 27 Brady 27] was shot in the head in 1981 by gunman [28 28 John Hinckley 28] , who tried to kill [29 29 Reagan 29] in a deranged bid to impress [30 30 Jodie Foster 30] , an actress he never met but with whom he was obsessed . The [31 31 Brady 31] bill , calling for a waiting period before someone could buy a gun so a background check could be made , was first introduced in [32 32 Congress 32] in 1987 but it took seven years to pass because of opposition from the [33 33 National Rifle Associa tion 33] gun lobby .

1: 214323 James_Brady

2: 3356 Bill_Clinton

3: 6886 Chicago

4: 1850769 Sarah_Brady

5: 32070 Republican_Party_(United_States)

```
6: 25433 Ronald_Reagan
7: 3356 Bill_Clinton
8: 1850769 Sarah_Brady
9: 5043544 Democratic_Party_(United_States)
10: 31756 United_States_Congress
11: 3356 Bill_Clinton
12: 3434750 United_States
13: 5043192 Hillary_Rodham_Clinton
14: 28504 San_Diego
15: 32070 Republican_Party_(United_States)
16: 25433 Ronald_Reagan
17: 5043544 Democratic_Party_(United_States)
18: 32070 Republican_Party_(United_States)
19: 32070 Republican_Party_(United_States)
20: 1850769 Sarah_Brady
21: 3356 Bill_Clinton
22: 3356 Bill_Clinton
23: 22658078 Remington_Model_8
24: 1348 AK-47
25: 83239 Uzi
26: 214323 James_Brady
27: 214323 James_Brady
28: 222294 John_Hinckley,_Jr.
29: 25433 Ronald_Reagan
30: 18717883 Jodie_Foster
31: 214323 James_Brady
32: 31756 United_States_Congress
33: 70101 National_Rifle_Association
1: 214323 James_Brady
    [winner: 528805 James_(band), pem_pos=0,
final_scores=-0.34848917, final_local_scores=-0.2865924,
pem=-1.6834015, similarity=4.6918235]
    RECALL MISS: 214323 James_Brady]
1: 214323 James_Brady
    [pem_pos=2, final_scores=0.5012401,
final_local_scores=0.5437437, pem=-0.49995595,
similarity=23.034948]
2: 3356 Bill_Clinton
    [pem_pos=1, final_scores=0.8070918,
final_local_scores=0.7866196, pem=0.0, similarity=27.56401]
3: 6886 Chicago
    [pem_pos=0, final_scores=0.6261433,
final_local_scores=0.6612311, pem=-0.20658566,
similarity=24.945614]
4: 1850769 Sarah_Brady
    [pem_pos=0, final_scores=0.50142074,
final_local_scores=0.57178915, pem=0.0, similarity=21.153042]
5: 32070 Republican_Party_(United_States)
    [pem_pos=1, final_scores=0.44096535,
final_local_scores=0.42194772, pem=-0.1310191,
similarity=17.393976]
6: 25433 Ronald_Reagan
```

```
[pem_pos=0, final_scores=0.42964345,
final_local_scores=0.37011898, pem=0.0, similarity=15.134802]
7: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.5384419,
final_local_scores=0.49112087, pem=0.0, similarity=18.745739]
8: 1850769 Sarah_Brady
    [pem_pos=0, final_scores=0.332677,
final_local_scores=0.38746232, pem=0.0, similarity=15.652361]
9: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.50091594,
final_local_scores=0.48923707, pem=-0.076093495,
similarity=19.103333]
10: 31756 United_States_Congress
    [pem_pos=1, final_scores=0.45900863,
final_local_scores=0.43668646, pem=-0.3349529,
similarity=18.942835]
11: 3356 Bill_Clinton
    [pem_pos=1, final_scores=0.8568345,
final_local_scores=0.840956, pem=0.0, similarity=29.185516]
12: 3434750 United_States
    [pem_pos=0, final_scores=0.48461694,
final_local_scores=0.5091028, pem=-0.073457584,
similarity=19.681831]
13: 5043192 Hillary_Rodham_Clinton
    [pem_pos=0, final_scores=0.5504872,
final_local_scores=0.53260773, pem=0.0, similarity=19.98379]
14: 28504 San_Diego
    [pem_pos=0, final_scores=0.41013402,
final_local_scores=0.45652872, pem=-0.18777876,
similarity=18.73461]
15: 32070 Republican_Party_(United_States)
    [pem_pos=1, final_scores=0.42903483,
final_local_scores=0.40891546, pem=-0.1310191,
similarity=17.005068]
16: 25433 Ronald_Reagan
    [pem_pos=0, final_scores=0.40261477,
final_local_scores=0.34249842, pem=0.0, similarity=14.310549]
17: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.4482494,
final_local_scores=0.43170685, pem=-0.076093495,
similarity=17.386515]
18: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.44938642,
final_local_scores=0.4311465, pem=-0.1310191,
similarity=17.668488]
19: 32070 Republican_Party_(United_States)
    [pem_pos=1, final_scores=0.42678475,
final_local_scores=0.4064576, pem=-0.1310191,
similarity=16.93172]
20: 1850769 Sarah_Brady
    [pem_pos=0, final_scores=0.3178212,
final_local_scores=0.3712346, pem=0.0, similarity=15.168094]
```

```
21: 3356 Bill_Clinton
    [pem_pos=1, final_scores=0.8831299,
final_local_scores=0.86967975, pem=0.0, similarity=30.042692]
22: 3356 Bill_Clinton
    [pem_pos=1, final_scores=0.85760707,
final_local_scores=0.8417999, pem=0.0, similarity=29.210701]
23: 22658078 Remington_Model_8
    [pem_pos=13, final_scores=0.24394475,
final_local_scores=0.3887071, pem=-3.2865312,
similarity=33.562176]
24: 1348 AK-47
    [pem_pos=0, final_scores=0.5065858,
final_local_scores=0.63668275, pem=0.0, similarity=23.089596]
25: 83239 Uzi
    [pem_pos=0, final_scores=0.38254386,
final_local_scores=0.519423, pem=0.0, similarity=19.590332]
26: 214323 James_Brady
    [pem_pos=1, final_scores=0.4287176,
final_local_scores=0.4645239, pem=-0.49995595,
similarity=20.670872]
27: 214323 James_Brady
    [pem_pos=2, final_scores=0.40345204,
final_local_scores=0.4369251, pem=-0.49995595,
similarity=19.847267]
28: 222294 John_Hinckley,_Jr.
    [pem_pos=0, final_scores=0.4611264,
final_local_scores=0.50357616, pem=0.0, similarity=19.11743]
29: 25433 Ronald_Reagan
    [pem_pos=0, final_scores=0.47118354,
final_local_scores=0.41739935, pem=0.0, similarity=16.545742]
30: 18717883 Jodie_Foster
    [pem_pos=0, final_scores=0.49899155,
final_local_scores=0.562594, pem=0.0, similarity=20.87864]
31: 214323 James_Brady
    [pem_pos=2, final_scores=0.32717294,
final_local_scores=0.35360175, pem=-0.49995595,
similarity=17.360733]
32: 31756 United_States_Congress
    [pem_pos=0, final_scores=0.40930778,
final_local_scores=0.38239574, pem=-0.3349529,
similarity=17.322691]
33: 70101 National_Rifle_Association
    [pem_pos=0, final_scores=0.4390549,
final_local_scores=0.43876714, pem=0.0, similarity=17.1834]
```

Sample: b'618_Latest'

Latest opinion polls on [1 1 German 1] political parties . [2 2 BONN 2] 1996-08-27 Here are the latest opinion polls tracking national support for [3 3 Germany 3] 's main political parties : AUGUST 1996 CDU / CSU [4 4 SPD 4] [5 5 FDP 5] [6 6 Greens 6] PDS Emnid Aug 25 41.0 34.0 7.0 10.0 6.0 Elect Res Aug 23 41.0 35.0 5.0 11.0 4.0 [7 7 Allensbach 7]

Aug 21 37.2 32.8 8.0 13.0 5.6 Emnid Aug 18 41.0 34.0 6.0 10.0
5.0 JULY 1996 CDU / CSU [8 8 SPD 8] [9 9 FDP 9] [10 10 Greens
10] PDS Emnid July 7 39.0 32.0 7.0 11.0 5.0 Elect Res July
40.0 33.0 6.0 12.0 4.0 JUNE 1996 CDU / CSU [11 11 SPD 11] [12
12 FDP 12] [13 13 Greens 13] PDS Emnid June 30 39.0 33.0 6.0
12.0 5.0 Elect Res June 21 42.0 33.0 6.0 12.0 4.0 [14 14 Alle
nsbach 14] June 12 37.4 32.8 7.3 12.3 5.4 [15 15 Forsa 15]
June 6 39.0 36.0 6.0 12.0 5.0 MAY 1996 CDU / CSU [16 16 SPD 16
] [17 17 FDP 17] [18 18 Greens 18] PDS Emnid May 26 40.0 31.0
6.0 13.0 6.0 Elect Res May 25 43.0 32.0 6.0 12.0 4.0 [19 19 Fo
rsa 19] May 23 38.0 37.0 7.0 11.0 5.0 [20 20 Allensbach 20]
May 15 38.5 32.5 8.1 12.0 4.4 APRIL 1996 CDU / CSU [21 21 SPD
21] [22 22 FDP 22] [23 23 Greens 23] PDS Emnid April 28 40.0
32.0 5.0 11.0 5.0 Elect Res April 20 43.0 32.0 6.0 12.0 4.0 [2
4 24 Allensbach 24] April 17 38.1 32.3 6.5 12.9 6.3 OFFICIAL
RESULTS OF THE OCTOBER 16 , 1994 GENERAL ELECTION : CDU / CSU
[25 25 SPD 25] [26 26 FDP 26] [27 27 Greens 27] [28 28 PDS 28]
41.5 36.4 6.9 7.3 4.4 NOTE : Elect Res = Electoral Research
Group (Forschungsgruppe Wahlen) -- [29 29 Bonn 29]
newsroom , +49 228 2609760

1: 11867 Germany
2: 3295 Bonn
3: 11867 Germany
4: 60585 Social_Democratic_Party_of_Germany
5: 10825 Free_Democratic_Party_(Germany)
6: 12246 Alliance_90/The_Greens
7: 10038362 Allensbach
8: 60585 Social_Democratic_Party_of_Germany
9: 10825 Free_Democratic_Party_(Germany)
10: 12246 Alliance_90/The_Greens
11: 60585 Social_Democratic_Party_of_Germany
12: 10825 Free_Democratic_Party_(Germany)
13: 12246 Alliance_90/The_Greens
14: 10038362 Allensbach
15: 19902641 Forsa_Institute
16: 60585 Social_Democratic_Party_of_Germany
17: 10825 Free_Democratic_Party_(Germany)
18: 12246 Alliance_90/The_Greens
19: 19902641 Forsa_Institute
20: 10038362 Allensbach
21: 60585 Social_Democratic_Party_of_Germany
22: 10825 Free_Democratic_Party_(Germany)
23: 12246 Alliance_90/The_Greens
24: 10038362 Allensbach
25: 60585 Social_Democratic_Party_of_Germany
26: 10825 Free_Democratic_Party_(Germany)
27: 12246 Alliance_90/The_Greens
28: 23326 Party_of_Democratic_Socialism_(Germany)
29: 3295 Bonn
1: 11867 Germany
[pem_pos=1, final_scores=0.77181584,
final_local_scores=0.81985897, pem=-1.0465823,

```
similarity=34.247417]
2: 3295 Bonn
    [pem_pos=0, final_scores=0.48160243,
final_local_scores=0.4860099, pem=0.0, similarity=18.593218]
3: 11867 Germany
    [pem_pos=0, final_scores=0.7054211,
final_local_scores=0.74733275, pem=-0.15529114,
similarity=27.23611]
4: 60585 Social_Democratic_Party_of_Germany
    [pem_pos=0, final_scores=0.47542578,
final_local_scores=0.41619867, pem=-0.22288324,
similarity=17.721985]
5: 10825 Free_Democratic_Party_(Germany)
    [pem_pos=0, final_scores=0.5993039,
final_local_scores=0.5626768, pem=-0.41945136,
similarity=23.162151]
6: 12246 Alliance_ '90/The_Greens
    [pem_pos=1, final_scores=0.64489335,
final_local_scores=0.6066118, pem=-1.9227095,
similarity=32.648212]
7: 10038362 Allensbach
    [pem_pos=0, final_scores=0.506365,
final_local_scores=0.6151319, pem=0.0, similarity=22.446476]
8: 60585 Social_Democratic_Party_of_Germany
    [pem_pos=0, final_scores=0.4853751,
final_local_scores=0.4270668, pem=-0.22288324,
similarity=18.046312]
9: 10825 Free_Democratic_Party_(Germany)
    [pem_pos=0, final_scores=0.59308845,
final_local_scores=0.55588734, pem=-0.41945136,
similarity=22.959541]
10: 12246 Alliance_ '90/The_Greens
    [pem_pos=1, final_scores=0.65278596,
final_local_scores=0.61523324, pem=-1.9227095,
similarity=32.905495]
11: 60585 Social_Democratic_Party_of_Germany
    [pem_pos=0, final_scores=0.46814126,
final_local_scores=0.40824145, pem=-0.22288324,
similarity=17.484526]
12: 10825 Free_Democratic_Party_(Germany)
    [pem_pos=0, final_scores=0.5866603,
final_local_scores=0.54886556, pem=-0.41945136,
similarity=22.749998]
13: 12246 Alliance_ '90/The_Greens
    [pem_pos=1, final_scores=0.6626104,
final_local_scores=0.62596494, pem=-1.9227095,
similarity=33.22575]
14: 10038362 Allensbach
    [pem_pos=0, final_scores=0.5858911,
final_local_scores=0.7020021, pem=0.0, similarity=25.038857]
15: 19902641 Forsa_Institute
    [pem_pos=0, final_scores=0.53256816,
```

```
final_local_scores=0.6092858, pem=0.0, similarity=22.272017]
16: 60585 Social_Democratic_Party_of_Germany
    [pem_pos=0, final_scores=0.47066766,
final_local_scores=0.41100115, pem=-0.22288324,
similarity=17.566881]
17: 10825 Free_Democratic_Party_(Germany)
    [pem_pos=0, final_scores=0.58552116,
final_local_scores=0.5476213, pem=-0.41945136,
similarity=22.712866]
18: 12246 Alliance_90/The_Greens
    [pem_pos=1, final_scores=0.6738585,
final_local_scores=0.63825184, pem=-1.9227095,
similarity=33.592415]
19: 19902641 Forsa_Institute
    [pem_pos=0, final_scores=0.46134776,
final_local_scores=0.5314883, pem=0.0, similarity=19.950384]
20: 10038362 Allensbach
    [pem_pos=0, final_scores=0.58207464,
final_local_scores=0.6978332, pem=0.0, similarity=24.914448]
21: 60585 Social_Democratic_Party_of_Germany
    [pem_pos=0, final_scores=0.49319762,
final_local_scores=0.43561172, pem=-0.22288324,
similarity=18.30131]
22: 10825 Free_Democratic_Party_(Germany)
    [pem_pos=0, final_scores=0.59369475,
final_local_scores=0.5565496, pem=-0.41945136,
similarity=22.979305]
23: 12246 Alliance_90/The_Greens
    [pem_pos=1, final_scores=0.6652431,
final_local_scores=0.62884074, pem=-1.9227095,
similarity=33.31157]
24: 10038362 Allensbach
    [pem_pos=0, final_scores=0.59576935,
final_local_scores=0.7127926, pem=0.0, similarity=25.360865]
25: 60585 Social_Democratic_Party_of_Germany
    [pem_pos=0, final_scores=0.5169232,
final_local_scores=0.4615283, pem=-0.22288324,
similarity=19.074713]
26: 10825 Free_Democratic_Party_(Germany)
    [pem_pos=0, final_scores=0.54421383,
final_local_scores=0.5024993, pem=-0.41945136,
similarity=21.366335]
27: 12246 Alliance_90/The_Greens
    [pem_pos=1, final_scores=0.6401319,
final_local_scores=0.60141057, pem=-1.9227095,
similarity=32.492996]
28: 23326 Party_of_Democratic_Socialism_(Germany)
    [pem_pos=0, final_scores=0.2912051,
final_local_scores=0.26971048, pem=-1.1049166,
similarity=18.14712]
29: 3295 Bonn
    [pem_pos=0, final_scores=0.492557,
```

```
final_local_scores=0.49797606, pem=0.0, similarity=18.950312]
```

Sample: b'619_Most'

Most hostages freed from hijacked [1 1 Sudanese 1] plane . [2 2 STANSTED 2] , [3 3 England 3] 1996-08-27 Armed hijackers believed to be [4 4 Iraqis 4] released 140 people on Tuesday from a [5 5 Sudan Airways 5] plane carrying 199 passengers and crew that landed in [6 6 London 6] after being diverted on a flight from [7 7 Khartoum 7] to [8 8 Amman 8] , police said . Police spokesman Roger Grimwade said the six or seven hijackers remained on board the aircraft , which arrived from [9 9 Cyprus 9] at 4.30 a.m. (0330 [10 10 GMT 10]) . He said they had made various requests to police negotiators but stepped back from earlier suggestions that the hijackers had asked to speak to a [11 11 British-based 11] [12 12 Iraqi 12] police named as Mr Sadiki . The hijackers , who have said they want to seek asylum in [13 13 Britain 13] , are believed to be armed with grenades and possibly other explosives , according to police . They earlier threatened to blow up the aircraft . The Airbus A310 was diverted first to [14 14 Cyprus 14] , then to [15 15 Britain 15] .

1: 27421 Sudan

2: 65456 London_Stansted_Airport

3: 9316 England

4: 7515928 Iraq

5: 383994 Sudan_Airways

6: 17867 London

7: 17320 Khartoum

8: 45007 Amman

9: 5593 Cyprus

10: 12701 Greenwich_Mean_Time

11: 31717 United_Kingdom

12: 7515928 Iraq

13: 31717 United_Kingdom

14: 5593 Cyprus

15: 31717 United_Kingdom

1: 27421 Sudan

[pem_pos=0, final_scores=0.5295241,
final_local_scores=0.51568943, pem=-0.19893567,
similarity=20.560757]

2: 65456 London_Stansted_Airport

[pem_pos=0, final_scores=0.30802804,
final_local_scores=0.3568279, pem=-0.2670293,
similarity=16.190315]

3: 9316 England

[pem_pos=0, final_scores=0.6304106,
final_local_scores=0.6429106, pem=-0.19053856,
similarity=24.311626]

4: 7515928 Iraq

[pem_pos=1, final_scores=0.49453205,
final_local_scores=0.47124094, pem=-0.8129937,
similarity=22.57367]

```
5: 383994 Sudan_Airways
    [pem_pos=0, final_scores=0.50189936,
final_local_scores=0.59349245, pem=0.0, similarity=21.80071]
6: 17867 London
    [pem_pos=0, final_scores=0.46281666,
final_local_scores=0.44297463, pem=-0.09086737,
similarity=17.80311]
7: 17320 Khartoum
    [pem_pos=0, final_scores=0.4359904,
final_local_scores=0.41886365, pem=-0.03966028,
similarity=16.805119]
8: 45007 Amman
    [pem_pos=0, final_scores=0.43974692,
final_local_scores=0.42285174, pem=0.0, similarity=16.708452]
9: 5593 Cyprus
    [pem_pos=0, final_scores=0.5898574,
final_local_scores=0.5755759, pem=-0.3752463,
similarity=23.306692]
10: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.6465216,
final_local_scores=0.71864676, pem=0.0, similarity=25.535566]
11: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.71358806,
final_local_scores=0.6938938, pem=0.0, similarity=24.796888]
12: 7515928 Iraq
    [pem_pos=0, final_scores=0.56637996,
final_local_scores=0.54972386, pem=-0.48469657,
similarity=23.130424]
13: 31717 United_Kingdom
    [pem_pos=1, final_scores=0.5317746,
final_local_scores=0.49529022, pem=-0.61114657,
similarity=22.193672]
14: 5593 Cyprus
    [pem_pos=0, final_scores=0.5371737,
final_local_scores=0.51802695, pem=-0.3752463,
similarity=21.589317]
15: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.54658425,
final_local_scores=0.5114675, pem=-0.61114657,
similarity=22.676434]
```

Sample: b'620_British'

British Data in merger talks with Mentmore . [1 1 LONDON 1]
1996-08-27 Mentmore Abbey said on Tuesday that merger
discussions were taking place with the board of British Data
Management , an information resource management and archive
storage company . But it noted in a brief statement : "
However , it is too early to say at this stage whether or not
terms can be agreed . " Mentmore Abbey , the UK stationery
and housewares business formerly known as Platignum , said it
was making the statement in response to recent newspaper
articles linking the two groups . Shares in Mentmore Abbey

edged two pence higher to 81.5 pence , valuing the group at just under 30 million stg . British Data Management 's shares slipped 0.5 pence to 179.5p , valuing that company at around 45 million stg . -- London Newsroom +44 171 542 4017

1: 17867 London

1: 17867 London

[pem_pos=0, final_scores=0.43472707,
final_local_scores=0.7474605, pem=-0.09086737,
similarity=26.889574]

Sample: b'621_Hijacked'

Hijacked [1 1 Sudan 1] plane lands at a [2 2 London 2] airport . [3 3 STANSTED 3] , [4 4 England 4] 1996-08-27 A hijacked [5 5 Sudan Airways 5] plane carrying 199 passengers and crew landed at [6 6 Stansted 6] airport just outside [7 7 London 7] early on Tuesday morning after flying from [8 8 Cyprus 8] , eyewitnesses said . The Airbus 310 Flight 150 , which was hijacked on Monday evening on its way from [9 9 Khartoum 9] to the [10 10 Jordanian 10] capital [11 11 Amman 11] , landed at 4.30 a.m. (0330 [12 12 GMT 12]) after a flight of more than four hours . At least one of the unidentified hijackers was apparently armed with grenades and TNT and threatened to blow the plane up in [13 13 Cyprus 13] unless it was refuelled so they could fly to [14 14 Britain 14] to claim political asylum . " They will surrender the passengers there and surrender themselves , " police spokesman Glafcos Xenos told reporters at [15 15 Cyprus 15] 's [16 16 Larnaca 16] airport . Security was tight at [17 17 Stansted 17] airport , which is about 30 miles (50 km) northeast of the capital . [18 18 Stansted 18] has been designated as the preferred option for handling hijackings in southern [19 19 England 19] because it is more remote than [20 20 Heathrow 20] and [21 21 Gatwick 21] , [22 22 London 22] 's two major airports , and handles less air traffic .

1: 27421 Sudan

2: 17867 London

3: 65456 London_Stansted_Airport

4: 9316 England

5: 383994 Sudan_Airways

6: 65456 London_Stansted_Airport

7: 17867 London

8: 5593 Cyprus

9: 17320 Khartoum

10: 7515964 Jordan

11: 45007 Amman

12: 12701 Greenwich_Mean_Time

13: 5593 Cyprus

14: 31717 United_Kingdom

15: 5593 Cyprus

16: 287852 Larnaca

17: 65456 London_Stansted_Airport

18: 65456 London_Stansted_Airport

```
19: 9316 England
20: 13595 London_Heathrow_Airport
21: 101359 Gatwick_Airport
22: 17867 London
1: 27421 Sudan
    [pem_pos=0, final_scores=0.5772267,
final_local_scores=0.595175, pem=-0.2848764,
similarity=23.400126]
2: 17867 London
    [pem_pos=0, final_scores=0.5724916,
final_local_scores=0.5509727, pem=-0.09086737,
similarity=21.025988]
3: 65456 London_Stansted_Airport
    [pem_pos=0, final_scores=0.41683668,
final_local_scores=0.41641116, pem=-0.2670293,
similarity=17.9684]
4: 9316 England
    [pem_pos=0, final_scores=0.6047746,
final_local_scores=0.6091391, pem=-0.19053856,
similarity=23.303816]
5: 383994 Sudan_Airways
    [pem_pos=0, final_scores=0.5120395,
final_local_scores=0.5916883, pem=0.0, similarity=21.746872]
6: 65456 London_Stansted_Airport
    [pem_pos=0, final_scores=0.43890584,
final_local_scores=0.44051838, pem=-0.2670293,
similarity=18.687807]
7: 17867 London
    [pem_pos=0, final_scores=0.50988543,
final_local_scores=0.48258495, pem=-0.09086737,
similarity=18.98516]
8: 5593 Cyprus
    [pem_pos=0, final_scores=0.5248259,
final_local_scores=0.5187115, pem=-0.3752463,
similarity=21.609747]
9: 17320 Khartoum
    [pem_pos=0, final_scores=0.39945358,
final_local_scores=0.38597596, pem=-0.03966028,
similarity=15.823685]
10: 7515964 Jordan
    [pem_pos=0, final_scores=0.45813143,
final_local_scores=0.45062482, pem=-0.18608934,
similarity=18.549238]
11: 45007 Amman
    [pem_pos=0, final_scores=0.44427437,
final_local_scores=0.43460095, pem=0.0, similarity=17.059072]
12: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.67810583,
final_local_scores=0.74481, pem=0.0, similarity=26.316328]
13: 5593 Cyprus
    [pem_pos=0, final_scores=0.5779739,
final_local_scores=0.5767677, pem=-0.3752463,
```

```
similarity=23.342258]
14: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.51760066,
final_local_scores=0.49693578, pem=-0.61114657,
similarity=22.242779]
15: 5593 Cyprus
    [pem_pos=0, final_scores=0.58675915,
final_local_scores=0.5863642, pem=-0.3752463,
similarity=23.628637]
16: 287852 Larnaca
    [pem_pos=0, final_scores=0.39305848,
final_local_scores=0.41837853, pem=0.0, similarity=16.574963]
17: 65456 London_Stansted_Airport
    [pem_pos=0, final_scores=0.4886877,
final_local_scores=0.49489748, pem=-0.2670293,
similarity=20.310587]
18: 65456 London_Stansted_Airport
    [pem_pos=0, final_scores=0.41620487,
final_local_scores=0.415721, pem=-0.2670293,
similarity=17.947803]
19: 9316 England
    [pem_pos=0, final_scores=0.5209093,
final_local_scores=0.517529, pem=-0.19053856,
similarity=20.569988]
20: 13595 London_Heathrow_Airport
    [pem_pos=0, final_scores=0.43133324,
final_local_scores=0.41776174, pem=0.0, similarity=16.556557]
21: 101359 Gatwick_Airport
    [pem_pos=0, final_scores=0.37907702,
final_local_scores=0.37705672, pem=0.0, similarity=15.341838]
22: 17867 London
    [pem_pos=0, final_scores=0.5622743,
final_local_scores=0.5398118, pem=-0.09086737,
similarity=20.692924]
```

Sample: b'622_Hijacked'

Hijacked [1 1 Sudan 1] plane expected at [2 2 London 2] 's [3 3 Stansted 3] . [4 4 LONDON 4] 1996-08-27 A hijacked [5 5 Sud an Airways 5] plane with 199 passengers and crew on board was expected to land at [6 6 London 6] 's [7 7 Stansted 7] airport later on Tuesday morning , a police spokeswoman said . " That is the plan at the moment . That is where the plane is being directed to , " Ruth Collin of [8 8 Essex 8] police , the force responsible for [9 9 Stansted 9] , said . [10 10 Stansted 10] , [11 11 London 11] 's third-busiest airport after [12 12 Heathrow 12] and [13 13 Gatwick 13] , is located about 30 miles (48 km) northeast of the capital . [1 4 14 British 14] officials said they would much prefer to deal with a hijacking at [15 15 Stansted 15] because of its relatively remote location and because air traffic would be less badly disrupted there than at [16 16 Heathrow 16] or [17 17 Gatwick 17] . They said police and the emergency services

were implementing a well-rehearsed contingency plan to handle the hijacking . The armed hijackers of the Airbus 310 Flight 150 , which is expected to arrive about 4 a.m. (0300 [18 18 G MT 18]) , have said they intend to surrender and seek political asylum in [19 19 Britain 19] . The plane was hijacked on its way from [20 20 Khartoum 20] to the [21 21 Jordanian 21] capital [22 22 Amman 22] on Monday evening and landed at [23 23 Larnaca 23] airport in [24 24 Cyprus 24] to refuel . The identity and number of the hijackers was not known . One of them negotiated through the pilot in [25 25 English 25] . The pilot said several hijackers appeared to be placed around the plane .

1: 27421 Sudan
2: 17867 London
3: 65456 London_Stansted_Airport
4: 17867 London
5: 383994 Sudan_Airways
6: 17867 London
7: 65456 London_Stansted_Airport
8: 40100758 Essex
9: 65456 London_Stansted_Airport
10: 65456 London_Stansted_Airport
11: 17867 London
12: 13595 London_Heathrow_Airport
13: 101359 Gatwick_Airport
14: 31717 United_Kingdom
15: 65456 London_Stansted_Airport
16: 13595 London_Heathrow_Airport
17: 101359 Gatwick_Airport
18: 12701 Greenwich_Mean_Time
19: 31717 United_Kingdom
20: 17320 Khartoum
21: 7515964 Jordan
22: 45007 Amman
23: 287852 Larnaca
24: 5593 Cyprus
25: 8569916 English_language
1: 27421 Sudan
[pem_pos=0, final_scores=0.49477118,
final_local_scores=0.520758, pem=-0.2848764,
similarity=21.179373]
2: 17867 London
[pem_pos=0, final_scores=0.65196353,
final_local_scores=0.63810265, pem=-0.09086737,
similarity=23.62612]
3: 65456 London_Stansted_Airport
[pem_pos=0, final_scores=0.48392886,
final_local_scores=0.47013998, pem=-0.2670293,
similarity=19.571774]
4: 17867 London
[pem_pos=0, final_scores=0.675064,
final_local_scores=0.6633364, pem=-0.09086737,

```
similarity=24.379145]
5: 383994 Sudan_Airways
    [pem_pos=0, final_scores=0.5242212,
final_local_scores=0.59904015, pem=0.0, similarity=21.966267]
6: 17867 London
    [pem_pos=0, final_scores=0.6128389,
final_local_scores=0.59536487, pem=-0.09086737,
similarity=22.350739]
7: 65456 London_Stansted_Airport
    [pem_pos=0, final_scores=0.49258655,
final_local_scores=0.47959727, pem=-0.2670293,
similarity=19.853998]
8: 40100758 Essex
    [pem_pos=0, final_scores=0.37267786,
final_local_scores=0.37359995, pem=-0.576716,
similarity=18.374952]
9: 65456 London_Stansted_Airport
    [pem_pos=0, final_scores=0.4673496,
final_local_scores=0.45202965, pem=-0.2670293,
similarity=19.031326]
10: 65456 London_Stansted_Airport
    [pem_pos=0, final_scores=0.4625497,
final_local_scores=0.44678652, pem=-0.2670293,
similarity=18.87486]
11: 17867 London
    [pem_pos=0, final_scores=0.4907382,
final_local_scores=0.46198845, pem=-0.09086737,
similarity=18.37052]
12: 13595 London_Heathrow_Airport
    [pem_pos=0, final_scores=0.47511095,
final_local_scores=0.44539726, pem=0.0, similarity=17.381256]
13: 101359 Gatwick_Airport
    [pem_pos=0, final_scores=0.4216882,
final_local_scores=0.40245724, pem=0.0, similarity=16.09984]
14: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5065737,
final_local_scores=0.48861033, pem=-0.48936236,
similarity=21.332048]
15: 65456 London_Stansted_Airport
    [pem_pos=0, final_scores=0.45213884,
final_local_scores=0.43541425, pem=-0.2670293,
similarity=18.53549]
16: 13595 London_Heathrow_Airport
    [pem_pos=0, final_scores=0.3329484,
final_local_scores=0.29010627, pem=0.0, similarity=12.747063]
17: 101359 Gatwick_Airport
    [pem_pos=0, final_scores=0.3080786,
final_local_scores=0.278356, pem=0.0, similarity=12.396411]
18: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.72658974,
final_local_scores=0.793686, pem=0.0, similarity=27.774885]
19: 31717 United_Kingdom
```

```
[pem_pos=0, final_scores=0.50845337,
final_local_scores=0.4906636, pem=-0.61114657,
similarity=22.055603]
20: 17320 Khartoum
    [pem_pos=0, final_scores=0.39933366,
final_local_scores=0.39018548, pem=-0.03966028,
similarity=15.949305]
21: 7515964 Jordan
    [pem_pos=0, final_scores=0.4619642,
final_local_scores=0.4734152, pem=-0.18608934,
similarity=19.22935]
22: 45007 Amman
    [pem_pos=0, final_scores=0.46799582,
final_local_scores=0.46995217, pem=0.0, similarity=18.114023]
23: 287852 Larnaca
    [pem_pos=0, final_scores=0.34876293,
final_local_scores=0.38681573, pem=0.0, similarity=15.633066]
24: 5593 Cyprus
    [pem_pos=0, final_scores=0.5293043,
final_local_scores=0.5506493, pem=-0.3752463,
similarity=22.562834]
25: 8569916 English_language
    [pem_pos=0, final_scores=0.49171692,
final_local_scores=0.5572086, pem=-0.640266,
similarity=24.199795]
```

Sample: b'623_Painted'

Painted parrot scam lands [1 1 Australian 1] in jail . [2 2 P ERT 2] , [3 3 Australia 3] 1996-08-27 A conman who painted common [4 4 Australian 4] parrots with dye to make them look like rare birds worth thousands of dollars was jailed for fraud on Tuesday . Denham Peiris painted six green parrots , worth about A\$ 100 ([5 5 US\$ 5] 79) a pair , with a cinnamon hair dye and traded them as Indian Ringneck Parrots , valued at A\$ 14,000 a pair . Peiris , 32 , of [6 6 P erth 6] , was sentenced in the [7 7 Western Australian 7] District Court to two years in jail for fraud after trading three pairs of impostor birds for 21 parrots worth a total of A\$ 28,000 . " Everyone was fooled , " said pet shop owner Shane Drew , who unknowingly traded the disguised birds . " I 'd already had three local breeders have a look at them and sent photos and videos of the birds to the eastern states for authentication -- they said they 're nice birds , " Drew told reporters outside court . " After I was told they were dyed , I checked them over again . It was a perfect paint job except for one feather under a wing of one bird that was only half dyed , " Drew said . Drew said Peiris , a bird enthusiast , would have succeeded with the fraud if he had not tried to trade a fourth pair of bogus birds after an associate had told police of the scam . If not for the informant , the painted parrots ' true colours would not have been known for six months , he said . " They moult in the summer , so five

or six months down the track , I would have looked like the guilty party , " Drew said .

1: 4689264 Australia
2: 24355 Perth
3: 4689264 Australia
4: 4689264 Australia
5: 18717338 United_States_dollar
6: 24355 Perth
7: 33613 Western_Australia
1: 4689264 Australia
 [pem_pos=0, final_scores=0.70760334,
final_local_scores=0.6633708, pem=-0.491009,
similarity=26.556204]
2: 24355 Perth
 [pem_pos=0, final_scores=0.517695,
final_local_scores=0.47783315, pem=-0.32334754,
similarity=20.10762]
3: 4689264 Australia
 [pem_pos=0, final_scores=0.78698987,
final_local_scores=0.7500885, pem=-0.17168477,
similarity=27.4075]
4: 4689264 Australia
 [pem_pos=0, final_scores=0.6940325,
final_local_scores=0.6485467, pem=-0.491009,
similarity=26.113823]
5: 18717338 United_States_dollar
 [pem_pos=0, final_scores=0.4561268,
final_local_scores=0.5395507, pem=0.0, similarity=20.190983]
6: 24355 Perth
 [pem_pos=0, final_scores=0.45383006,
final_local_scores=0.4080704, pem=-0.32334754,
similarity=18.02576]
7: 33613 Western_Australia
 [pem_pos=0, final_scores=0.5869068,
final_local_scores=0.56490016, pem=0.0, similarity=20.94746]

Sample: b'624_NZ'

NZ motorist 's arrest brings free flight to [1 1 Tonga 1] . [2 2 WELLINGTON 2] 1996-08-27 A [3 3 New Zealand 3] motorist got an unexpected free flight to [4 4 Tonga 4] on Tuesday after being caught drinking and driving . The man drew attention to himself in the [5 5 North Island 5] town of [6 6 Tauranga 6] while trying to reverse his car out of a pothole on Saturday night . He spun the wheels so much that the tyres caught alight and smoke began pouring from under the bonnet , the [7 7 New Zealand Press Association 7] reported . Police arrested the man and charged him with drink-driving , but then discovered he was wanted by the Immigration Service as an overstayer . The man was due to catch a flight to Nuku'alofa in [8 8 Tonga 8] later on Tuesday .

1: 30158 Tonga
2: 33804 Wellington

```
3: 4913064 New_Zealand
4: 30158 Tonga
5: 62046 North_Island
6: 31495916 Tauranga
7: 3185516 New_Zealand_Press_Association
8: 30158 Tonga
1: 30158 Tonga
    [pem_pos=0, final_scores=0.5373904,
final_local_scores=0.5070281, pem=-0.4741999,
similarity=21.799215]
2: 33804 Wellington
    [pem_pos=0, final_scores=0.49148422,
final_local_scores=0.47123545, pem=-0.9265605,
similarity=23.1911]
3: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.50242025,
final_local_scores=0.4762379, pem=-0.23795462,
similarity=19.595636]
4: 30158 Tonga
    [pem_pos=0, final_scores=0.47432363,
final_local_scores=0.43813717, pem=-0.4741999,
similarity=19.743372]
5: 62046 North_Island
    [pem_pos=0, final_scores=0.50867134,
final_local_scores=0.5186888, pem=0.0, similarity=19.56842]
6: 31495916 Tauranga
    [pem_pos=0, final_scores=0.32466674,
final_local_scores=0.31704363, pem=0.0, similarity=13.550927]
7: 3185516 New_Zealand_Press_Association
    [pem_pos=0, final_scores=0.42968023,
final_local_scores=0.55644444, pem=0.0, similarity=20.695124]
8: 30158 Tonga
    [pem_pos=0, final_scores=0.48976892,
final_local_scores=0.4550088, pem=-0.4741999,
similarity=20.246857]
```

```
Sample: b'625_Indonesia'
[1 1 Indonesia 1] 's SBPUs auction results . [2 2 JAKARTA 2]
1996-08-27 The following is the result of central bank
securities ( SBPUs ) auction on Tuesday at 0800 [3 3 GMT 3] :
SBPUs seven-day 14-day Cut-off-rate ( percent ) 15.75 16.00
Total ( in billion rupiah ) 38.43 218.50 -- [4 4 Jakarta 4]
newsroom +6221 384-6364
1: 14579 Indonesia
2: 16275 Jakarta
3: 12701 Greenwich_Mean_Time
4: 16275 Jakarta
1: 14579 Indonesia
    [pem_pos=0, final_scores=0.5855828,
final_local_scores=0.5912351, pem=-0.23151328,
similarity=22.992353]
2: 16275 Jakarta
```

```
[pem_pos=0, final_scores=0.4390589,
final_local_scores=0.3894776, pem=0.0, similarity=15.7125025]
3: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.6367618,
final_local_scores=0.7556703, pem=0.0, similarity=26.640423]
4: 16275 Jakarta
    [pem_pos=0, final_scores=0.5101143,
final_local_scores=0.46709484, pem=0.0, similarity=18.028755]
```

Sample: b'626_Seoul'

[1 1 Seoul 1] embassies warned of terrorist attacks . [2 2 SE
OUL 2] 1996-08-27 [3 3 South Korea 3] has told its overseas
missions to step up security since its embassy in the former [
4 4 Yugoslavia 4] received a threat over [5 5 Seoul 5] 's
crackdown on radical students , the foreign ministry said on
Tuesday . A letter to the [6 6 Belgrade 6] embassy on Monday
under the name of the Macedonian Communist Party demanded [7 7
South Korea 7] release detained student leaders , a ministry
spokesman said . " The letter said the party would assault
the embassy , other South Korean-related facilities and [8 8 K
orean 8] nationals unless our authorities released arrested
students , " he said . Nearly 400 members of an outlawed
student group were arrested after violent protests demanding
reunification with communist [9 9 North Korea 9] were crushed
by riot police at a [10 10 Seoul 10] university this month .
Authorities branded the violence , in which a police officer
was killed , as an act of terror . " The ministry has ordered
the embassy to take urgent security measures against possible
terrorist attacks , " the official said , asking not to be
named . It had called for similar precautions at other
overseas missions , including those in [11 11 Canada 11] and [
12 12 Bangladesh 12] where leftist groups have staged
protests over the crackdown .

```
1: 19159283 Seoul
2: 19159283 Seoul
3: 27019 South_Korea
4: 34244 Yugoslavia
5: 19159283 Seoul
6: 55904 Belgrade
7: 27019 South_Korea
8: 27019 South_Korea
9: 21255 North_Korea
10: 19159283 Seoul
11: 5042916 Canada
12: 3454 Bangladesh
1: 19159283 Seoul
    [pem_pos=0, final_scores=0.45671993,
final_local_scores=0.399033, pem=-0.0378014,
similarity=16.203224]
2: 19159283 Seoul
    [pem_pos=0, final_scores=0.48456782,
final_local_scores=0.4294526, pem=-0.0378014,
```

```
similarity=17.111006]
3: 27019 South_Korea
    [pem_pos=0, final_scores=0.44873643,
final_local_scores=0.38386965, pem=-0.22449642,
similarity=16.765995]
4: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.38042372,
final_local_scores=0.37910503, pem=-0.6536551,
similarity=18.957642]
5: 19159283 Seoul
    [pem_pos=0, final_scores=0.5306396,
final_local_scores=0.47977906, pem=-0.0378014,
similarity=18.612848]
6: 55904 Belgrade
    [pem_pos=0, final_scores=0.35596806,
final_local_scores=0.3593226, pem=-0.05193398,
similarity=15.095042]
7: 27019 South_Korea
    [pem_pos=0, final_scores=0.49189943,
final_local_scores=0.43101865, pem=-0.22449642,
similarity=18.173016]
8: 27019 South_Korea
    [pem_pos=3, final_scores=0.2907002,
final_local_scores=0.21205941, pem=-2.69666,
similarity=25.082846]
9: 21255 North_Korea
    [pem_pos=0, final_scores=0.4782477,
final_local_scores=0.45472693, pem=-0.3090239,
similarity=19.34019]
10: 19159283 Seoul
    [pem_pos=0, final_scores=0.55333084,
final_local_scores=0.5045658, pem=-0.0378014,
similarity=19.352533]
11: 5042916 Canada
    [pem_pos=0, final_scores=0.5785312,
final_local_scores=0.59763193, pem=-0.11662793,
similarity=22.558483]
12: 3454 Bangladesh
    [pem_pos=0, final_scores=0.48235852,
final_local_scores=0.5157356, pem=-0.28971618,
similarity=21.055815]
```

Sample: b'627_Nuclear'

Nuclear pact will be a step to disarmament-China . [1 1 BEIJING 1] 1996-08-27 [2 2 China 2] on Tuesday reaffirmed its support for a global nuclear test ban treaty blocked by [3 3 India 3] last week , saying the pact would be an important step in achieving total nuclear disarmament . " Although the final draft of the treaty probably did n't totally satisfy any country , it was in general balanced , " the official People 's Daily newspaper said in a commentary . [4 4 India 4] blocked the [5 5 Comprehensive Test Ban Treaty 5] ([6 6 CTBT

[6]) at the Conference on Disarmament in [7 7 Geneva 7] , saying the pact did not contain a clause committing the five declared nuclear powers to a timetable for nuclear disarmament . [8 8 New Delhi 8] 's stance , which was seen as effectively thwarting 2-1/2 years of negotiations at the Conference on Disarmament , drew widespread but generally muted foreign criticism . [9 9 India 9] has also pledged to oppose any forwarding of the draft treaty to the United Nations General Assembly . [10 10 China 10] said many countries had compromised to complete the treaty and that the issue of a disarmament schedule could be discussed in future negotiations . " The completion of the test ban treaty would be an important and practical step in the gradual process of achieving total nuclear disarmament , " it said . [11 11 China 11] has pledged support for the pact since reaching a deal with the [12 12 United States 12] that made international inspections of nuclear sites more difficult than in earlier drafts of the accord . The newspaper said [13 13 China 13] advocated the complete ban and destruction of nuclear weapons but that there was little hope other nuclear powers would soon adopt the same stance . " Some nuclear powers stubbornly uphold policies of nuclear deterrence based on first use of nuclear weapons , " it said . [14 14 China 14] held on July 29 what it said would be its last nuclear test before a self-imposed moratorium that took effect the following day . [15 15 China 15] was the last declared nuclear power to announce a halt to testing .

1: 18603746 Beijing
2: 5405 China
3: 14533 India
4: 14533 India
5: 7851 Comprehensive_Nuclear-Test-Ban_Treaty
6: 7851 Comprehensive_Nuclear-Test-Ban_Treaty
7: 12521 Geneva
8: 51585 New_Delhi
9: 14533 India
10: 5405 China
11: 5405 China
12: 3434750 United_States
13: 5405 China
14: 5405 China
15: 5405 China
1: 18603746 Beijing
[pem_pos=0, final_scores=0.5019837,
final_local_scores=0.5059747, pem=-0.042161036,
similarity=19.418287]
2: 5405 China
[pem_pos=0, final_scores=0.55821365,
final_local_scores=0.50530523, pem=-0.15669827,
similarity=20.02118]
3: 14533 India
[pem_pos=0, final_scores=0.6478235,

```
final_local_scores=0.6274303, pem=-0.13622136,
similarity=23.554277]
4: 14533 India
    [pem_pos=0, final_scores=0.64120954,
final_local_scores=0.6202056, pem=-0.13622136,
similarity=23.338678]
5: 7851 Comprehensive_Nuclear-Test-Ban_Treaty
    [pem_pos=0, final_scores=0.48534137,
final_local_scores=0.5949603, pem=0.0, similarity=21.844515]
6: 7851 Comprehensive_Nuclear-Test-Ban_Treaty
    [pem_pos=0, final_scores=0.7199022,
final_local_scores=0.8511824, pem=0.0, similarity=29.490694]
7: 12521 Geneva
    [pem_pos=0, final_scores=0.47182685,
final_local_scores=0.5071308, pem=-0.24772356,
similarity=20.570667]
8: 51585 New_Delhi
    [pem_pos=0, final_scores=0.42988783,
final_local_scores=0.4436937, pem=0.0, similarity=17.330418]
9: 14533 India
    [pem_pos=0, final_scores=0.5892698,
final_local_scores=0.5634693, pem=-0.13622136,
similarity=21.645554]
10: 5405 China
    [pem_pos=0, final_scores=0.4769526,
final_local_scores=0.41653997, pem=-0.15669827,
similarity=17.372246]
11: 5405 China
    [pem_pos=0, final_scores=0.51680255,
final_local_scores=0.46006995, pem=-0.15669827,
similarity=18.671268]
12: 3434750 United_States
    [pem_pos=0, final_scores=0.5024473,
final_local_scores=0.4921453, pem=-0.073457584,
similarity=19.175785]
13: 5405 China
    [pem_pos=0, final_scores=0.62883836,
final_local_scores=0.58245206, pem=-0.15669827,
similarity=22.323395]
14: 5405 China
    [pem_pos=0, final_scores=0.5130571,
final_local_scores=0.45597863, pem=-0.15669827,
similarity=18.549175]
15: 5405 China
    [pem_pos=0, final_scores=0.5048169,
final_local_scores=0.44697744, pem=-0.15669827,
similarity=18.280561]
```

Sample: b'628_Japan'

[1 1 Japan 1] declares food poisoning threat receding . [2 2 TOKYO 2] 1996-08-27 The government said on Tuesday that the threat from a mysterious killer germ , while still requiring

vigilance , appears to be receding in the western [3 3 Japan 3] city of [4 4 Sakai 4] , where the epidemic hit the hardest . The food poisoning epidemic caused by the O-157 colon bacillus in [5 5 Sakai 5] appears to be " settling down " , Health Minister [6 6 Naoto Kan 6] was quoted by a government spokesman as telling the cabinet . The O-157 colon bacillus has been found responsible for a widespread food poisoning epidemic that has killed 11 people and made over 9,500 ill this year . [7 7 Sakai 7] , near the regional commercial centre of [8 8 Osaka 8] , has been hit hardest by the deadly bacteria , with nearly 6,500 , mostly schoolchildren , affected by the disease . Two children in [9 9 Sakai 9] have died from complications associated with the bacteria . [10 10 Kan 10] told a cabinet meeting on Tuesday that no new victims have been reported since August 8 , indicating that the peak has passed , at least for [11 11 Sakai 11] . [12 12 Sakai 12] officials agreed with the assessment , but said it was too early to feel relieved . " There are still patients hospitalised and problems which must be dealt with , " a city spokesman said , citing the issue of whether to allow children infected with the bacteria but not showing symptoms to attend school from September . Health authorities believe school lunches were the source of the food poisoning in [13 13 Sakai 13] , but researchers have been unable to pinpoint the exact source of the infection . The outbreak has prompted authorities to tighten sanitary standards at slaughterhouses and meatpacking plants and sparked calls for an overhaul of the nation 's school lunch programme . The ministers agreed at Tuesday 's cabinet meeting to step up inspection measures for school lunches in September and October , when schools around the country resume . [14 14 Japan 14] 's Agriculture Ministry also announced that it will compile hygiene guidelines based on [15 15 U.S. 15] government methods of checking the safety of farm produce to prevent another outbreak of the epidemic . As of Monday , 31 children were still hospitalised in [16 16 Sakai 16] city , of whom six were in serious condition .

- 1: 15573 Japan
- 2: 30057 Tokyo
- 3: 15573 Japan
- 4: 466743 Sakai,_Osaka
- 5: 466743 Sakai,_Osaka
- 6: 365175 Naoto_Kan
- 7: 466743 Sakai,_Osaka
- 8: 52230 Osaka
- 9: 466743 Sakai,_Osaka
- 10: 365175 Naoto_Kan
- 11: 466743 Sakai,_Osaka
- 12: 466743 Sakai,_Osaka
- 13: 466743 Sakai,_Osaka
- 14: 15573 Japan
- 15: 3434750 United_States

```
16: 466743 Sakai,_Osaka
1: 15573 Japan
    [pem_pos=0, final_scores=0.6219098,
final_local_scores=0.62824124, pem=-0.13549899,
similarity=23.574549]
2: 30057 Tokyo
    [pem_pos=0, final_scores=0.6490972,
final_local_scores=0.66029733, pem=-0.09283106,
similarity=24.299131]
3: 15573 Japan
    [pem_pos=0, final_scores=0.78267676,
final_local_scores=0.8038547, pem=-0.13549899,
similarity=28.815205]
4: 466743 Sakai,_Osaka
    [pem_pos=0, final_scores=0.48528558,
final_local_scores=0.4333558, pem=-0.6650015,
similarity=20.638296]
5: 466743 Sakai,_Osaka
    [pem_pos=0, final_scores=0.55468816,
final_local_scores=0.50916755, pem=-0.6650015,
similarity=22.90067]
6: 365175 Naoto_Kan
    [pem_pos=0, final_scores=0.52229166,
final_local_scores=0.5763254, pem=0.0, similarity=21.288412]
7: 466743 Sakai,_Osaka
    [pem_pos=0, final_scores=0.51701814,
final_local_scores=0.46801883, pem=-0.6650015,
similarity=21.672709]
8: 52230 Osaka
    [pem_pos=0, final_scores=0.42348033,
final_local_scores=0.3954954, pem=-0.08992924,
similarity=16.381134]
9: 466743 Sakai,_Osaka
    [pem_pos=0, final_scores=0.5301598,
final_local_scores=0.48237407, pem=-0.6650015,
similarity=22.101099]
10: 365175 Naoto_Kan
    [pem_pos=0, final_scores=0.37822258,
final_local_scores=0.41895187, pem=-1.114921,
similarity=22.655186]
11: 466743 Sakai,_Osaka
    [pem_pos=0, final_scores=0.5259352,
final_local_scores=0.4777593, pem=-0.6650015,
similarity=21.963385]
12: 466743 Sakai,_Osaka
    [pem_pos=0, final_scores=0.5260009,
final_local_scores=0.47783113, pem=-0.6650015,
similarity=21.965528]
13: 466743 Sakai,_Osaka
    [pem_pos=0, final_scores=0.5117392,
final_local_scores=0.46225232, pem=-0.6650015,
similarity=21.500626]
```

```
14: 15573 Japan
    [pem_pos=0, final_scores=0.65843385,
final_local_scores=0.6681382, pem=-0.13549899,
similarity=24.765154]
15: 3434750 United_States
    [pem_pos=0, final_scores=0.42006963,
final_local_scores=0.4869061, pem=-0.14075686,
similarity=19.38542]
16: 466743 Sakai,_Osaka
    [pem_pos=0, final_scores=0.54813164,
final_local_scores=0.5020056, pem=-0.6650015,
similarity=22.686941]
```

Sample: b'629_Japan'

[1 1 Japan 1] aluminium shipments surge 8.9 pct in July . [2 2 TOKYO 2] 1996-08-27 [3 3 Japanese 3] shipments of aluminium mill products in July surged 8.9 percent over the same month last year to 224,609 tonnes , while production rose 6.4 percent to 222,457 tonnes , according to preliminary data released on Tuesday by the Japan Aluminium Federation . The surge reflected strong demand from the beverage can and housing construction sectors , federation officials said . Figures released two weeks ago from [4 4 Japan 4] 's seven largest mills already showed that [5 5 Japan 5] 's can sheet output reached its highest monthly level ever in July , reflecting above-average temperatures that sparked a jump in beer consumption . A federation official added that the half-year through September 1996 also appeared likely to set a new record for can production , despite cooler temperatures in August . July inventories stood at 75,632 tonnes , down 2.6 percent from the prior month . July foil output was off 0.2 percent year-on-year at 11,525 tonnes , while foil shipments fell 0.8 percent to 11,244 tonnes . June mill output data were revised slightly downward to 210,622 tonnes from the preliminary 210,683 , while shipments were revised marginally upward to 213,989 tonnes from 213,845 . Final figures for June pegged can stock shipments at 40,144 tonnes , up 2.4 percent year-on-year . Shipments to the auto sector fell 5.5 percent to 15,286 tonnes , while the construction sector rose 2.4 percent to 79,390 tonnes . Exports dipped 3.7 percent to 18,867 tonnes . -- Tokyo Commodities Desk (81-3 3432 6179)

```
1: 15573 Japan
2: 30057 Tokyo
3: 15573 Japan
4: 15573 Japan
5: 15573 Japan
1: 15573 Japan
    [pem_pos=0, final_scores=0.70100224,
final_local_scores=0.642388, pem=-0.13549899,
similarity=23.996716]
2: 30057 Tokyo
    [pem_pos=0, final_scores=0.6973633,
```

```
final_local_scores=0.7110945, pem=-0.09283106,
similarity=25.815022]
3: 15573 Japan
    [pem_pos=0, final_scores=0.8786176,
final_local_scores=0.8364057, pem=-0.99076414,
similarity=34.437656]
4: 15573 Japan
    [pem_pos=0, final_scores=0.7573636,
final_local_scores=0.7039543, pem=-0.13549899,
similarity=25.833977]
5: 15573 Japan
    [pem_pos=0, final_scores=0.69990313,
final_local_scores=0.64118737, pem=-0.13549899,
similarity=23.960886]
```

Sample: b'630_China'

[1 1 China 1] thanks [2 2 Gabon 2] for support on human rights . [3 3 BEIJING 3] 1996-08-27 [4 4 China 4] has publicly thanked [5 5 Gabon 5] for its strong support at the [6 6 United Nations Human Rights Commission 6] , where [7 7 Beijing 7] has come under attack from [8 8 Western 8] nations for its human rights record . [9 9 China 9] 's President [10 10 Jiang Zemin 10] offered his gratitude in a meeting on Monday with visiting [11 11 Gabon 11] President [12 12 Omar Bongo 12] , the official [13 13 Xinhua 13] news agency said . [14 14 Jiang 14] also acknowledged the [15 15 Central African 15] nation 's support for [16 16 China 16] 's stance on [17 17 Taiwan 17] , which [18 18 Beijing 18] views as a renegade province , while [19 19 Bongo 19] was quoted as thanking [20 20 China 20] for economic and technological aid , the agency said late on Monday . In April , [21 21 China 21] quashed a draft resolution by the [22 22 U.N. Human Rights Commission 1] expressing concern over continuing reports of [23 23 Beijing 23] 's violations of fundamental freedoms . After the defeat of the resolution , drafted by the [24 24 European Union 24] and the [25 25 United States 25] , [26 26 China 26] thanked 26 countries for backing its motion for " no action " on the document . It was the sixth year in a row that [27 27 China 27] avoided censure at the [28 28 U.N. 28] 's main human rights forum . [29 29 China 29] has been accused of a wide range of human rights abuses , often in violation of its own legal code , in an effort to silence dissent . A [30 30 Xinhua 30] commentary earlier this year said a plot by the [31 31 West 31] to force its human rights standards and values on other countries was doomed to failure .

1: 5405 China
2: 12027 Gabon
3: 18603746 Beijing
4: 5405 China
5: 12027 Gabon
6: 50883 United_Nations_Commission_on_Human_Rights

```
7: 18603746 Beijing
8: 21208200 Western_world
9: 5405 China
10: 95871 Jiang_Zemin
11: 12027 Gabon
12: 392403 Omar_Bongo
13: 263168 Xinhua_News_Agency
14: 95871 Jiang_Zemin
15: 5478 Central_African_Republic
16: 5405 China
17: 25734 Taiwan
18: 18603746 Beijing
19: 392403 Omar_Bongo
20: 5405 China
21: 5405 China
22: 18603746 Beijing
23: 9317 European_Union
24: 3434750 United_States
25: 5405 China
26: 317153 Foreign_minister
27: 5405 China
28: 31769 United_Nations
29: 5405 China
30: 263168 Xinhua_News_Agency
31: 21208200 Western_world
1: 5405 China
    [pem_pos=0, final_scores=0.5418275,
final_local_scores=0.4968341, pem=-0.15669827,
similarity=19.768383]
2: 12027 Gabon
    [pem_pos=0, final_scores=0.48390663,
final_local_scores=0.5087548, pem=-0.3021721,
similarity=20.91523]
3: 18603746 Beijing
    [pem_pos=0, final_scores=0.49446124,
final_local_scores=0.48896027, pem=-0.042161036,
similarity=18.910542]
4: 5405 China
    [pem_pos=0, final_scores=0.52713567,
final_local_scores=0.4807855, pem=-0.15669827,
similarity=19.289461]
5: 12027 Gabon
    [pem_pos=0, final_scores=0.47911686,
final_local_scores=0.5035227, pem=-0.3021721,
similarity=20.759094]
6: 50883 United_Nations_Commission_on_Human_Rights
    [pem_pos=0, final_scores=0.36270565,
final_local_scores=0.40296578, pem=0.0, similarity=16.115015]
7: 18603746 Beijing
    [pem_pos=0, final_scores=0.36242265,
final_local_scores=0.34472817, pem=-0.042161036,
similarity=14.606367]
```

```
8: 21208200 Western_world
    [pem_pos=0, final_scores=0.38691896,
final_local_scores=0.39366257, pem=-1.5917408,
similarity=24.493523]
9: 5405 China
    [pem_pos=0, final_scores=0.525446,
final_local_scores=0.47893977, pem=-0.15669827,
similarity=19.23438]
10: 95871 Jiang_Zemin
    [pem_pos=0, final_scores=0.58232695,
final_local_scores=0.61113733, pem=0.0, similarity=22.32727]
11: 12027 Gabon
    [pem_pos=0, final_scores=0.5820395,
final_local_scores=0.61595, pem=-0.3021721,
similarity=24.11415]
12: 392403 Omar_Bongo
    [pem_pos=0, final_scores=0.7722514,
final_local_scores=0.8494505, pem=0.0, similarity=29.43901]
13: 263168 Xinhua_News_Agency
    [pem_pos=0, final_scores=0.4686821,
final_local_scores=0.48597497, pem=0.0, similarity=18.592175]
14: 95871 Jiang_Zemin
    [pem_pos=0, final_scores=0.53405267,
final_local_scores=0.5654507, pem=0.0, similarity=20.96389]
15: 5478 Central_African_Republic
    [pem_pos=0, final_scores=0.24079911,
final_local_scores=0.23337772, pem=-0.070826046,
similarity=11.439333]
16: 5405 China
    [pem_pos=0, final_scores=0.60732645,
final_local_scores=0.5683818, pem=-0.15669827,
similarity=21.90351]
17: 25734 Taiwan
    [pem_pos=0, final_scores=0.46782786,
final_local_scores=0.43466473, pem=-0.077886656,
similarity=17.484535]
18: 18603746 Beijing
    [pem_pos=0, final_scores=0.43496352,
final_local_scores=0.42396802, pem=-0.042161036,
similarity=16.971043]
19: 392403 Omar_Bongo
    [pem_pos=0, final_scores=0.7646342,
final_local_scores=0.8498137, pem=0.0, similarity=29.449848]
20: 5405 China
    [pem_pos=0, final_scores=0.67011744,
final_local_scores=0.6369713, pem=-0.15669827,
similarity=23.950357]
21: 5405 China
    [pem_pos=0, final_scores=0.59523255,
final_local_scores=0.555171, pem=-0.15669827,
similarity=21.509274]
22: 18603746 Beijing
```

```
[pem_pos=0, final_scores=0.4267007,
final_local_scores=0.4149422, pem=-0.042161036,
similarity=16.701694]
23: 9317 European_Union
    [pem_pos=0, final_scores=0.51544887,
final_local_scores=0.53590417, pem=-0.005814796,
similarity=20.113785]
24: 3434750 United_States
    [pem_pos=0, final_scores=0.53887105,
final_local_scores=0.54341817, pem=-0.073457584,
similarity=20.70587]
25: 5405 China
    [pem_pos=0, final_scores=0.5593943,
final_local_scores=0.5160231, pem=-0.15669827,
similarity=20.34102]
26: 317153 Foreign_minister
    [pem_pos=0, final_scores=0.35712093,
final_local_scores=0.3709433, pem=-0.35490593,
similarity=17.089436]
27: 5405 China
    [pem_pos=0, final_scores=0.58194834,
final_local_scores=0.54066, pem=-0.15669827,
similarity=21.076237]
28: 31769 United_Nations
    [pem_pos=0, final_scores=0.5458551,
final_local_scores=0.55329096, pem=0.0, similarity=20.601019]
29: 5405 China
    [pem_pos=0, final_scores=0.5249712,
final_local_scores=0.4784211, pem=-0.15669827,
similarity=19.218903]
30: 263168 Xinhua_News_Agency
    [pem_pos=0, final_scores=0.4184541,
final_local_scores=0.43110847, pem=0.0, similarity=16.95485]
31: 21208200 Western_world
    [pem_pos=0, final_scores=0.2998771,
final_local_scores=0.29858252, pem=-1.6371275,
similarity=21.902966]
1: 50883 United_Nations_Commission_on_Human_Rights
    [pem_pos=0, final_scores=0.049955532,
final_local_scores=0.06133376, pem=0.0, similarity=5.920035]
```

Sample: b'631_Taiwan'

```
[1 1 Taiwan 1] 's Cooperative Bank cuts prime lending rate .
[2 2 TAIPEI 2] 1996-08-27 Cooperative Bank of Taiwan , one of
the island 's leading state-run banks , said it was cutting
its prime lending rate by 0.05 percentage point , effective
on Tuesday . The bank said in a statement it was cutting its
prime lending rate to 7.35 percent from 7.40 percent . It
also would cut its three-month-to-three-year time deposit
rates by between 0.05 and 0.10 percentage points . The moves
made Cooperative the first major bank to cut rates in
response to a call on Friday from central bank governor Sheu
```

Yuan-dong . In the central bank 's latest bid to jumpstart [3 3 Taiwan 3] 's sluggish economy , Sheu ordered a reduction of up to a half-percentage point in banks ' reserve requirements and called on commercial banks to pass on the savings in the form of interest rate reductions . -- Taipei Newsroom

(5080815)

1: 25734 Taiwan
2: 57648 Taipei
3: 25734 Taiwan
1: 25734 Taiwan
[pem_pos=0, final_scores=0.5498149,
final_local_scores=0.4921561, pem=-0.077886656,
similarity=19.200191]
2: 57648 Taipei
[pem_pos=0, final_scores=0.39672858,
final_local_scores=0.3819785, pem=0.0, similarity=15.488714]
3: 25734 Taiwan
[pem_pos=0, final_scores=0.49018753,
final_local_scores=0.42702228, pem=-0.077886656,
similarity=17.25647]

Sample: b'632_Palestinians'

[1 1 Palestinians 1] to strike over [2 2 Jerusalem 2] demolition . [3 3 JERUSALEM 3] 1996-08-27 [4 4 Palestinian 4] leaders in [5 5 Jerusalem 5] called on Tuesday for a two-hour strike to protest what they called [6 6 Israel 6] 's war on Arab East Jerusalem after police demolished a building there . [7 7 Israeli 7] police demolished a 10 metre (yard) by 20 metre structure in [8 8 Jerusalem 8] 's [9 9 Old City 9] they said had been built with funding from the Palestinian self-rule Authority for use as a social club . Police used a huge crane to lift a bulldozer over the [10 10 Old City 10] 's walls to reach the building amidst narrow alleys , witnesses said . Bystanders at the scene said work had begun on the building in 1991 and it had yet to be completed when the [11 11 Israeli 11] police bulldozed it . The [12 12 Palestinian Authority 12] was set up under the 1993 PLO-Israel interim peace deal . " There has been a call for a general strike between one (1000 [13 13 GMT 13]) and three o'clock , " [14 14 Palestinian 14] lawmaker Ahmed Hashem Zighayer told [15 15 Reuters 15] . " This is a war that has been declared on us and we want our people to come and see the site where they declared the war . " Prime Minister [16 16 Benjamin Netanyahu 16] 's government , which took office in June , has said it will not allow the Authority , set up under a 1993 interim peace deal to control parts of the [17 17 Gaza Strip 17] and West Bank , to operate in [18 18 Jerusalem 18] . [19 19 Israel 19] 's previous government held the same position but in general turned a blind eye to [20 20 Palestinian Authority 20] activity in the city . Tuesday 's demolition came a day after PLO officials said they had bowed to [21 21 Netanyahu 21] 's demand they close offices in [22 22

Jerusalem 22] . They said two of the three offices [23 23 Israel 23] wanted closed had been shut . [24 24 Netanyahu 24] has made closure of the three offices a condition for resuming peace negotiations with the Palestine Liberation Organisation (PLO) . [25 25 Israel 25] captured Arab East Jerusalem in the 1967 [26 26 Middle East 26] war and annexed it . It says it will never cede any part of the city . [27 27 Palestinians 27] want [28 28 East Jerusalem 28] as capital of a future state . The city is up for negotiation at final peace talks which have yet to resume under [29 29 Netanyahu 29]] .

```
1: 241405 State_of_Palestine
2: 16043 Jerusalem
3: 16043 Jerusalem
4: 241405 State_of_Palestine
5: 16043 Jerusalem
6: 9282173 Israel
7: 9282173 Israel
8: 16043 Jerusalem
9: 302972 Old_City_(Jerusalem)
10: 302972 Old_City_(Jerusalem)
11: 9282173 Israel
12: 24093 Palestinian_National_Authority
13: 12701 Greenwich_Mean_Time
14: 241405 State_of_Palestine
15: 18998750 Reuters
16: 56472 Benjamin_Netanyahu
17: 12047 Gaza_Strip
18: 16043 Jerusalem
19: 9282173 Israel
20: 24093 Palestinian_National_Authority
21: 56472 Benjamin_Netanyahu
22: 16043 Jerusalem
23: 9282173 Israel
24: 56472 Benjamin_Netanyahu
25: 9282173 Israel
26: 19323 Middle_East
27: 241405 State_of_Palestine
28: 343508 East_Jerusalem
29: 56472 Benjamin_Netanyahu
1: 241405 State_of_Palestine
    [pem_pos=5, final_scores=0.53355175,
final_local_scores=0.52462614, pem=-1.7909594,
similarity=29.485119]
2: 16043 Jerusalem
    [pem_pos=0, final_scores=0.50204927,
final_local_scores=0.4618975, pem=-0.13239029,
similarity=18.593615]
3: 16043 Jerusalem
    [pem_pos=0, final_scores=0.54685456,
final_local_scores=0.5108404, pem=-0.13239029,
similarity=20.054169]
```

```
4: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.52649575,
final_local_scores=0.51691854, pem=-2.8508017,
similarity=35.018696]
5: 16043 Jerusalem
    [pem_pos=0, final_scores=0.40517175,
final_local_scores=0.3560735, pem=-0.13239029,
similarity=15.435615]
6: 9282173 Israel
    [pem_pos=0, final_scores=0.67486376,
final_local_scores=0.6396112, pem=-0.21279564,
similarity=24.334204]
7: 9282173 Israel
    [pem_pos=0, final_scores=0.5129916,
final_local_scores=0.46279043, pem=-0.34790626,
similarity=19.79227]
8: 16043 Jerusalem
    [pem_pos=0, final_scores=0.54592586,
final_local_scores=0.50982594, pem=-0.13239029,
similarity=20.023895]
9: 302972 Old_City_(Jerusalem)
    [pem_pos=0, final_scores=0.43428677,
final_local_scores=0.4368956, pem=-0.17957826,
similarity=18.104124]
10: 302972 Old_City_(Jerusalem)
    [pem_pos=0, final_scores=0.41119784,
final_local_scores=0.40957785, pem=-0.17957826,
similarity=17.288908]
11: 9282173 Israel
    [pem_pos=0, final_scores=0.5879145,
final_local_scores=0.54463226, pem=-0.34790626,
similarity=22.234592]
12: 24093 Palestinian_National_Authority
    [pem_pos=0, final_scores=0.44270885,
final_local_scores=0.4077052, pem=0.0, similarity=16.256449]
13: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.62169194,
final_local_scores=0.7397648, pem=0.0, similarity=26.16577]
14: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.7288849,
final_local_scores=0.7379979, pem=-2.8508017,
similarity=41.616146]
15: 18998750 Reuters
    [pem_pos=0, final_scores=0.37845045,
final_local_scores=0.40909255, pem=0.0, similarity=16.29785]
16: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.6441724,
final_local_scores=0.6286005, pem=0.0, similarity=22.848406]
17: 12047 Gaza_Strip
    [pem_pos=0, final_scores=0.38907754,
final_local_scores=0.3529491, pem=0.0, similarity=14.622418]
18: 16043 Jerusalem
```

```
[pem_pos=0, final_scores=0.5067925,
final_local_scores=0.46707875, pem=-0.13239029,
similarity=18.748234]
19: 9282173 Israel
    [pem_pos=0, final_scores=0.56302637,
final_local_scores=0.51744574, pem=-0.21279564,
similarity=20.688541]
20: 24093 Palestinian_National_Authority
    [pem_pos=0, final_scores=0.424411,
final_local_scores=0.38771755, pem=0.0, similarity=15.659977]
21: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.4322304,
final_local_scores=0.40084958, pem=0.0, similarity=16.051865]
22: 16043 Jerusalem
    [pem_pos=0, final_scores=0.56178814,
final_local_scores=0.5271531, pem=-0.13239029,
similarity=20.540972]
23: 9282173 Israel
    [pem_pos=0, final_scores=0.6033104,
final_local_scores=0.56144994, pem=-0.21279564,
similarity=22.001715]
24: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.47020066,
final_local_scores=0.4423263, pem=0.0, similarity=17.289612]
25: 9282173 Israel
    [pem_pos=0, final_scores=0.5332275,
final_local_scores=0.48489505, pem=-0.21279564,
similarity=19.717163]
26: 19323 Middle_East
    [pem_pos=0, final_scores=0.3933748,
final_local_scores=0.397017, pem=0.0, similarity=15.937492]
27: 241405 State_of_Palestine
    [pem_pos=5, final_scores=0.49184686,
final_local_scores=0.4790699, pem=-1.7909594,
similarity=28.12563]
28: 343508 East_Jerusalem
    [pem_pos=0, final_scores=0.41384214,
final_local_scores=0.38607502, pem=0.0, similarity=15.610962]
29: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.45022482,
final_local_scores=0.42050576, pem=0.0, similarity=16.638443]
```

Sample: b'633_SEC'

SEC adopts rules to improve investor access to best stock prices . [1 1 WASHINGTON 1] 1996-08-28 The Securities and Exchange Commission Wednesday approved rules designed to give stock market investors a better chance at getting the best price available for their orders . The so-called order handling rules , which were proposed last September , were approved in a 4-0 vote . The new rules will require specialists at stock exchanges and market makers on [2 2 Nasd aq 2] to allow customers to view price quotes from other

electronic trading systems that may not be readily available to them . The rules will also require that customer limit orders be displayed with prices better than those available in quotes publicly available at the time . Specialists and market makers are companies or individuals recognised as qualified to maintain an orderly market in a stock . In a limit order , investors specify the price at which they are willing to buy or sell , as opposed to a market order executed at prevailing prices . " These rules are intended to empower all investors , by allowing their orders to compete on a level playing field , and by providing the disclosure they need to make informed decisions , " said SEC Chairman [3 3 Arthur Levitt 3] . The SEC 's goal , he added , was to create " one system where one price could be available to everybody . " The director of the SEC 's division of market regulation , Richard Lindsey , said [4 4 Wall Street 4] firms will probably need to spend about \$ 7 million to carry out the improvements called for by the new rules . In proposing the rules last year , the SEC noted that while technology has improved , commonplace practices still existed that worked against investors ' best interests . It noted that customers whose orders were not displayed lost the chance of getting the best possible price available in the market . It also cited the potential problem of a two-tiered market in which market makers quote one price to public investors while quoting better prices in private systems , thus robbing investors without access to " hidden " quotes the benefit of the best available prices . Earlier this month , the SEC settled charges of alleged malpractices on [5 5 Nasdaq 5] when the parent of the [6 6 Nasdaq 6] market , the National Association of Securities Dealers , agreed to spend \$ 100 million over five years to upgrade its oversight of brokers ' trading practices . The SEC dropped a third proposal that would have allowed investors with market orders to trade at a better price if there were shifts in prices before their orders were executed . The agency said it would monitor effects of the two new rules before considering the " price improvement " proposal again . The agency also proposed rule changes that would require continuous " bid " and " ask " quotations from exchange specialists and market makers who trade more than 1 percent of a stock during any quarter on [7 7 Nasdaq 7] . The SEC 's limit order display rule will mean a major change for [8 8 Nasdaq 8] , where such orders have never been displayed . The president of the [9 9 Nasdaq 9] , [10 10 Alfred Berkeley 10] , was to hold a news conference Wednesday afternoon to elaborate on the new rules ' effects on the market , the second largest in the world .

1: 108956 Washington,_D.C.

2: 21559 NASDAQ

3: 902148 Arthur_Levitt

4: 37274 Wall_Street

5: 21559 NASDAQ

```
6: 21559 NASDAQ
7: 21559 NASDAQ
8: 21559 NASDAQ
9: 21559 NASDAQ
10: 4277016 Alfred_Berkeley
1: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.5063407,
final_local_scores=0.59848636, pem=-1.7496481,
similarity=31.464598]
2: 21559 NASDAQ
    [pem_pos=0, final_scores=0.70649976,
final_local_scores=0.66491795, pem=0.0, similarity=23.93219]
3: 902148 Arthur_Levitt
    [pem_pos=0, final_scores=0.44523567,
final_local_scores=0.5079132, pem=0.0, similarity=19.246855]
4: 37274 Wall_Street
    [pem_pos=0, final_scores=0.35541564,
final_local_scores=0.37385786, pem=-0.039525714,
similarity=15.461326]
5: 21559 NASDAQ
    [pem_pos=0, final_scores=0.6794719,
final_local_scores=0.63539416, pem=0.0, similarity=23.051142]
6: 21559 NASDAQ
    [pem_pos=0, final_scores=0.80516976,
final_local_scores=0.7727, pem=0.0, similarity=27.148623]
7: 21559 NASDAQ
    [pem_pos=0, final_scores=0.61926335,
final_local_scores=0.5696255, pem=0.0, similarity=21.088474]
8: 21559 NASDAQ
    [pem_pos=0, final_scores=0.70035154,
final_local_scores=0.65820193, pem=0.0, similarity=23.731771]
9: 21559 NASDAQ
    [pem_pos=0, final_scores=0.6693631,
final_local_scores=0.6243518, pem=0.0, similarity=22.721617]
10: 4277016 Alfred_Berkeley
    [pem_pos=0, final_scores=0.5110495,
final_local_scores=0.596666, pem=0.0, similarity=21.895416]
```

Sample: b'634_ITALIAN'

[1 1 ITALIAN 1] CABINET APPROVES TELEVISION DECREE . [2 2 ROME 2] 1996-08-28 The [3 3 Italian 3] cabinet on Wednesday granted a reprieve for media mogul [4 4 Silvio Berlusconi 4]'s [5 5 Mediaset 5] television empire with a decree extending the current legal framework for television stations until January 31 , 1997 . The decree plugs a legal void in which magistrates could have forced the former prime minister to take one of the three stations he controls off the air because of a court ruling that no single proprietor should be allowed to keep three channels .

```
1: 14532 Italy
2: 25458 Rome
3: 14532 Italy
```

```
4: 26909 Silvio_Berlusconi
5: 1579765 Mediaset
1: 14532 Italy
    [pem_pos=0, final_scores=0.6045863,
final_local_scores=0.5759182, pem=-0.75948656,
similarity=25.406466]
2: 25458 Rome
    [pem_pos=0, final_scores=0.5033069,
final_local_scores=0.49161398, pem=-0.2517334,
similarity=20.12942]
3: 14532 Italy
    [pem_pos=0, final_scores=0.81193286,
final_local_scores=0.8024128, pem=-0.75948656,
similarity=32.165516]
4: 26909 Silvio_Berlusconi
    [pem_pos=0, final_scores=0.56810266,
final_local_scores=0.59580636, pem=0.0, similarity=21.869762]
5: 1579765 Mediaset
    [pem_pos=0, final_scores=0.40408653,
final_local_scores=0.4510976, pem=0.0, similarity=17.551365]
```

Sample: b'635_Passengers'

Passengers injured in train collision in [1 1 Linz 1] . [2 2 LINZ 2] , [3 3 Austria 3] 1996-08-28 A passenger train collided with a locomotive at a main railway station in [4 4 Linz 4] on Wednesday and police said around 10 people were injured . [5 5 Austrian 5] television reported earlier that more than 20 had been hurt in the accident at the station in [6 6 Linz 6] , 300 km (180 miles) west of [7 7 Vienna 7] . " One locomotive was stationary and the passenger train collided into the back of it , " a police spokesman in [8 8 Linz 8] told [9 9 Reuters 9] by telephone . The express passenger train travelling from [10 10 Steyr 10] , southeast of [11 11 Linz 11] , with around 80 people on board , hit the back of a service locomotive used to shunt wagons into sidings . The police spokesman said he was not sure whether any more passengers were still trapped in the wreckage . Doctors and medical staff from a hospital across the road from the station were at the scene of the accident within minutes and were able to treat the injured quickly , he added . Greater damage was averted as the driver of the passenger train had time to apply his emergency brakes before the collision occurred , a state railways spokesman told [12 1 2 Austrian 12] news agency APA .

```
1: 55856 Linz
2: 55856 Linz
3: 26964606 Austria
4: 55856 Linz
5: 26964606 Austria
6: 55856 Linz
7: 55866 Vienna
8: 55856 Linz
```

```
9: 18998750 Reuters
10: 923444 Steyr
11: 55856 Linz
12: 26964606 Austria
1: 55856 Linz
    [pem_pos=0, final_scores=0.518274,
final_local_scores=0.452586, pem=-0.08759329,
similarity=18.072126]
2: 55856 Linz
    [pem_pos=0, final_scores=0.63515335,
final_local_scores=0.5802589, pem=-0.08759329,
similarity=21.882141]
3: 26964606 Austria
    [pem_pos=0, final_scores=0.62908965,
final_local_scores=0.59829295, pem=-0.25262418,
similarity=23.317778]
4: 55856 Linz
    [pem_pos=0, final_scores=0.46545476,
final_local_scores=0.39488894, pem=-0.08759329,
similarity=16.350332]
5: 26964606 Austria
    [pem_pos=0, final_scores=0.6005039,
final_local_scores=0.5670674, pem=-0.40674555,
similarity=23.22408]
6: 55856 Linz
    [pem_pos=0, final_scores=0.5465128,
final_local_scores=0.48343253, pem=-0.08759329,
similarity=18.99265]
7: 55866 Vienna
    [pem_pos=0, final_scores=0.6342311,
final_local_scores=0.59504455, pem=-0.12517071,
similarity=22.527727]
8: 55856 Linz
    [pem_pos=0, final_scores=0.44451106,
final_local_scores=0.37201113, pem=-0.08759329,
similarity=15.667612]
9: 18998750 Reuters
    [pem_pos=0, final_scores=0.2683555,
final_local_scores=0.36837536, pem=0.0, similarity=15.082768]
10: 923444 Steyr
    [pem_pos=0, final_scores=0.43007243,
final_local_scores=0.44755632, pem=-0.11404446,
similarity=18.065878]
11: 55856 Linz
    [pem_pos=0, final_scores=0.48960245,
final_local_scores=0.42126662, pem=-0.08759329,
similarity=17.137495]
12: 26964606 Austria
    [pem_pos=0, final_scores=0.66947585,
final_local_scores=0.6424088, pem=-0.40674555,
similarity=25.472418]
```

Sample: b'636_Saskatchewan'

[1 1 Saskatchewan Wheat Pool 1] eyes hog market . [2 2 WINNIPEG 2] 1996-08-28 [3 3 Canada 3] 's largest grain handling firm said Wednesday it expects to forge a partnership with hog farmers by 1997 with a view to expanding the company 's scope into pork production . " [4 4 Saskatchewan 4] is well positioned to take advantage of growing world markets for pork , " [5 5 Saskatchewan Wheat Pool 5] chief executive Don Loewen said in a company statement . SWP said it was analyzing potential partnerships with hog farmers and expected the first deal to be in place in 1997 . The end of [6 6 Canada 6] 's rail freight subsidy last year caused a shift in feed grain production to the eastern Prairie . Analysts said livestock production would likely shift to the eastern Prairie rather than feed grains being shipped to the western Prairie . SWP said it may develop pork production systems that provide farmers with large integrated units , and it may consider contracting programs between producers and packers which would operate within the existing hog market framework . Saskatchewan Pork International Marketing Group has an export monopoly but some [7 7 Saskatchewan 7] farmers earlier this month called for an open market in hogs . --
Gilbert Le Gras 204 947 3548

1: 8025738 Saskatchewan_Wheat_Pool
2: 100730 Winnipeg
3: 5042916 Canada
4: 26840 Saskatchewan
5: 8025738 Saskatchewan_Wheat_Pool
6: 5042916 Canada
7: 26840 Saskatchewan
1: 8025738 Saskatchewan_Wheat_Pool
 [pem_pos=0, final_scores=0.5203369,
final_local_scores=0.56212425, pem=0.0, similarity=20.864622]
2: 100730 Winnipeg
 [pem_pos=0, final_scores=0.4890235,
final_local_scores=0.45729542, pem=-0.123312846,
similarity=18.406914]
3: 5042916 Canada
 [pem_pos=0, final_scores=0.80091864,
final_local_scores=0.7999252, pem=-0.11662793,
similarity=28.595318]
4: 26840 Saskatchewan
 [pem_pos=0, final_scores=0.4170456,
final_local_scores=0.3735426, pem=-0.08728397,
similarity=15.711634]
5: 8025738 Saskatchewan_Wheat_Pool
 [pem_pos=0, final_scores=0.48978597,
final_local_scores=0.52875197, pem=0.0, similarity=19.868727]
6: 5042916 Canada
 [pem_pos=0, final_scores=0.707032,
final_local_scores=0.69736826, pem=-0.11662793,
similarity=25.534813]

```
7: 26840 Saskatchewan
    [pem_pos=0, final_scores=0.45797426,
final_local_scores=0.41825098, pem=-0.08728397,
similarity=17.04582]
```

Sample: b'637_ATHLETICS'

ATHLETICS - ROVERETO INTERNATIONAL MEETING RESULTS . [1 1 R0
VERETO 1] , [2 2 Italy 2] 1996-08-28 Leading results from an
international athletics meeting on Wednesday : Women 's long
jump : 1. Ludmila Ninova ([3 3 Austria 3]) 6.72 metres 2. [4
4 Heike Drechsler 4] ([5 5 Germany 5]) 6.65 3. [6 6 Fiona Ma
y 6] ([7 7 Italy 7]) 6.64 Men 's 110 metres hurdles : 1. [8
8 Emilio Valle 8] ([9 9 Cuba 9]) 13.42 seconds 2. Steve
Brown ([10 10 U.S. 10]) 13.45 3. [11 11 Andrea Giaconi 11]
([12 12 Italy 12]) 13.80 Women 's 100 metres : 1. [13 13 Ch
andra Sturrup 13] ([14 14 Bahamas 14]) 11.34 seconds 2. [15
15 Natalya Voronova 15] ([16 16 Russia 16]) 11.53 3. Gabi
Rokmeier ([17 17 Germany 17]) 11.61 Men 's javelin : 1. [18
18 Sergey Makarov 18] ([19 19 Russia 19]) 85.26 metres 2. [2
0 20 Tom Pukstys 20] ([21 21 U.S. 21]) 84.20 3. [22 22 Peter
Blank 22] ([23 23 Germany 23]) 81.64 Men 's 100 metres : 1.
[24 24 Osmond Ezinwa 24] ([25 25 Nigeria 25]) 10.13 seconds
2. [26 26 Davidson Ezinwa 26] ([27 27 Nigeria 27]) 10.18 3.
[28 28 Stefano Tilli 28] ([29 29 Italy 29]) 10.43 Men 's
400 metres : 1. [30 30 Davis Kamoga 30] ([31 31 Uganda 31])
45.15 seconds 2. [32 32 Marco Vaccari 32] ([33 33 Italy
33]) 46.16 3. [34 34 Kennedy Ochieng 34] ([35 35 Kenya
35]) 46.21 Women 's pole vault : 1. Mariacarla Bresciani ([3
6 36 Italy 36]) 3.85 metres 2. Andrea Muller ([37 37 Germany
37]) 3.75 3. Nastja Rysich ([38 38 germany 38]) 3.75 Women
's 800 metres : 1. [39 39 Ana Fidelia Quirot 39] ([40 40 Cuba
40]) 1 minute 58.98 seconds 2. [41 41 Letitia Vriesde 41] ([
42 42 Surinam 42]) 2:00.39 3. [43 43 Inez Turner 43] ([44 44
Jamaica 44]) 2:00.91 Men 's high jump : 1. Wolfgang Kreissig
([45 45 Germany 45]) 2.20 metres 2. Kostantin Matusevitch ([46
46 Israel 46]) 2.20 3. Michele Buiatti ([47 47 Italy 47]
2.15 Men 's 800 metres : 1. Robert Kibet ([48 48 Kenya 48]
1:45.24 2. Vincent Malakwen ([49 49 Kenya 49]) 1:45.62 3.
Philip Kibitok ([50 50 Kenya 50]) 1:46.09 Women 's
javelin : 1. Oksana Ovchinnikova ([51 51 Russia 51]) 58.94
metres 2. [52 52 Natalya Shikolenko 52] ([53 53 Belarus
53]) 57.44 3. [54 54 Silke Renk 54] ([55 55 Germany 55])
56.70 Women 's 400 metres hurdles : 1. Virna De Angeli ([56 5
6 Italy 56]) 55.66 2. [57 57 Natalya Torshina 57] ([58 58 Ka
zakhstan 58]) 55.99 3. Anna Knoroz ([59 59 Russia 59])
57.02 Men 's 400 metres hurdles : 1. Lauren Ottoz ([60 60 It
aly 60]) 49.16 2. Brian Bronson ([61 61 U.S. 61]) 49.67 3.
John Ridgeon ([62 62 Britain 62]) 49.83 Men 's 3,000
metres : 1. [63 63 Luke Kipkosgei 63] ([64 64 Kenya 64])
7:46.91 2. [65 65 Alessandro Lambruschini 65] ([66 66 Italy 6
6]) 7:47.78 3. Richard Kosgei ([67 67 Kenya 67]) 7:48.38
1: 1412053 Rovereto

2: 14532 Italy
3: 26964606 Austria
4: 2805926 Heike_Drechsler
5: 11867 Germany
6: 2668040 Fiona_May
7: 14532 Italy
8: 11849304 Emilio_Valle
9: 5042481 Cuba
10: 3434750 United_States
11: 11880368 Andrea_Giaconi
12: 14532 Italy
13: 2511007 Chandra_Sturrup
14: 3451 The_Bahamas
15: 3572926 Natalya_Pomoshchnikova-Voronova
16: 25391 Russia
17: 11867 Germany
18: 2450892 Sergey_Makarov
19: 25391 Russia
20: 14086432 Tom_Pukstys
21: 3434750 United_States
22: 7712972 Peter_Blank
23: 11867 Germany
24: 2736304 Osmond_Ezinwa
25: 21383 Nigeria
26: 2732799 Davidson_Ezinwa
27: 21383 Nigeria
28: 18283614 Stefano_Tilli
29: 14532 Italy
30: 2434413 Davis_Kamoga
31: 31816 Uganda
32: 8169661 Marco_Vaccari
33: 14532 Italy
34: 8216969 Kennedy_Ochieng
35: 188171 Kenya
36: 14532 Italy
37: 11867 Germany
38: 11867 Germany
39: 30863751 Ana_Fidelia_Quirot
40: 5042481 Cuba
41: 2556569 Letitia_Vriesde
42: 26828 Suriname
43: 2736349 Inez_Turner
44: 15660 Jamaica
45: 11867 Germany
46: 9282173 Israel
47: 14532 Italy
48: 188171 Kenya
49: 188171 Kenya
50: 188171 Kenya
51: 25391 Russia
52: 3451457 Natalya_Shikolenko
53: 3457 Belarus

54: 3910212 Silke_Renk
55: 11867 Germany
56: 14532 Italy
57: 7771507 Natalya_Torshina-Alimzhanova
58: 16642 Kazakhstan
59: 25391 Russia
60: 14532 Italy
61: 3434750 United_States
62: 31717 United_Kingdom
63: 7754964 Luke_Kipkosgei
64: 188171 Kenya
65: 6815464 Alessandro_Lambruschini
66: 14532 Italy
67: 188171 Kenya
1: 1412053 Rovereto
 [pem_pos=0, final_scores=0.41437513,
final_local_scores=0.48189932, pem=0.0, similarity=18.47055]
2: 14532 Italy
 [pem_pos=0, final_scores=0.7443945,
final_local_scores=0.7401365, pem=-0.20612453,
similarity=27.2978]
3: 26964606 Austria
 [pem_pos=0, final_scores=0.684416,
final_local_scores=0.67041516, pem=-0.25262418,
similarity=25.470049]
4: 2805926 Heike_Drechsler
 [pem_pos=0, final_scores=0.8665646,
final_local_scores=1.0324986, pem=0.0, similarity=34.901527]
5: 11867 Germany
 [pem_pos=0, final_scores=0.8725011,
final_local_scores=0.88177544, pem=-0.15529114,
similarity=31.248148]
6: 2668040 Fiona_May
 [pem_pos=0, final_scores=0.4897998,
final_local_scores=0.6031912, pem=0.0, similarity=22.090141]
7: 14532 Italy
 [pem_pos=0, final_scores=0.77623934,
final_local_scores=0.77492213, pem=-0.20612453,
similarity=28.335873]
8: 11849304 Emilio_Valle
 [pem_pos=0, final_scores=0.50950235,
final_local_scores=0.6844461, pem=0.0, similarity=24.51495]
9: 5042481 Cuba
 [pem_pos=0, final_scores=0.56082654,
final_local_scores=0.55667955, pem=-0.23907764,
similarity=22.002283]
10: 3434750 United_States
 [pem_pos=0, final_scores=0.48857993,
final_local_scores=0.4716664, pem=-0.14075686,
similarity=18.930637]
11: 11880368 Andrea_Giaconi
 [pem_pos=0, final_scores=0.5611505,

```
final_local_scores=0.678838, pem=0.0, similarity=24.347593]
12: 14532 Italy
    [pem_pos=0, final_scores=0.8035967,
final_local_scores=0.8048059, pem=-0.20612453,
similarity=29.227663]
13: 2511007 Chandra_Sturrup
    [pem_pos=0, final_scores=0.7615989,
final_local_scores=0.9314763, pem=0.0, similarity=31.886824]
14: 3451 The_Bahamas
    [pem_pos=0, final_scores=0.5156239,
final_local_scores=0.53434193, pem=-0.35113433,
similarity=21.945065]
15: 3572926 Natalya_Pomoshchnikova-Voronova
    [pem_pos=0, final_scores=0.35403472,
final_local_scores=0.48397833, pem=0.0, similarity=18.53259]
16: 25391 Russia
    [pem_pos=0, final_scores=0.8137174,
final_local_scores=0.8347525, pem=-0.20299491,
similarity=30.104311]
17: 11867 Germany
    [pem_pos=0, final_scores=0.86144537,
final_local_scores=0.8696988, pem=-0.15529114,
similarity=30.887758]
18: 2450892 Sergey_Makarov
    [pem_pos=0, final_scores=0.5302134,
final_local_scores=0.67374694, pem=0.0, similarity=24.195665]
19: 25391 Russia
    [pem_pos=0, final_scores=0.8544206,
final_local_scores=0.8792146, pem=-0.20299491,
similarity=31.431147]
20: 14086432 Tom_Pukstys
    [pem_pos=0, final_scores=0.44171488,
final_local_scores=0.5803604, pem=0.0, similarity=21.408825]
21: 3434750 United_States
    [pem_pos=0, final_scores=0.5163927,
final_local_scores=0.5020476, pem=-0.14075686,
similarity=19.837273]
22: 7712972 Peter_Blank
    [pem_pos=0, final_scores=0.40235192,
final_local_scores=0.59638476, pem=0.0, similarity=21.887024]
23: 11867 Germany
    [pem_pos=0, final_scores=0.85110825,
final_local_scores=0.8584071, pem=-0.15529114,
similarity=30.550789]
24: 2736304 Osmond_Ezinwa
    [pem_pos=0, final_scores=0.6157607,
final_local_scores=0.78001, pem=0.0, similarity=27.366766]
25: 21383 Nigeria
    [pem_pos=0, final_scores=0.60639924,
final_local_scores=0.62569326, pem=-0.31455088,
similarity=24.472225]
26: 2732799 Davidson_Ezinwa
```

[pem_pos=0, final_scores=0.689314,
final_local_scores=0.85274607, pem=0.0, similarity=29.537357]
27: 21383 Nigeria
[pem_pos=0, final_scores=0.63414675,
final_local_scores=0.65600324, pem=-0.31455088,
similarity=25.376736]
28: 18283614 Stefano_Tilli
[pem_pos=0, final_scores=0.6432322,
final_local_scores=0.80248684, pem=0.0, similarity=28.037521]
29: 14532 Italy
[pem_pos=0, final_scores=0.8033717,
final_local_scores=0.8045601, pem=-0.20612453,
similarity=29.22033]
30: 2434413 Davis_Kamoga
[pem_pos=0, final_scores=0.5756326,
final_local_scores=0.7209625, pem=0.0, similarity=25.604673]
31: 31816 Uganda
[pem_pos=0, final_scores=0.5972522,
final_local_scores=0.62421495, pem=-0.2971513,
similarity=24.333487]
32: 8169661 Marco_Vaccari
[pem_pos=0, final_scores=0.44356662,
final_local_scores=0.579132, pem=0.0, similarity=21.372168]
33: 14532 Italy
[pem_pos=0, final_scores=0.8031692,
final_local_scores=0.80433893, pem=-0.20612453,
similarity=29.213728]
34: 8216969 Kennedy_Ochieng
[pem_pos=0, final_scores=0.465836,
final_local_scores=0.6283786, pem=0.0, similarity=22.841785]
35: 188171 Kenya
[pem_pos=0, final_scores=0.63946253,
final_local_scores=0.663811, pem=-0.27389407,
similarity=25.388636]
36: 14532 Italy
[pem_pos=0, final_scores=0.7924387,
final_local_scores=0.79261744, pem=-0.20612453,
similarity=28.863935]
37: 11867 Germany
[pem_pos=0, final_scores=0.85016245,
final_local_scores=0.8573739, pem=-0.15529114,
similarity=30.519957]
38: 11867 Germany
[pem_pos=0, final_scores=0.5326129,
final_local_scores=0.5104993, pem=-0.15529114,
similarity=20.168528]
39: 30863751 Ana_Fidelia_Quirot
[pem_pos=0, final_scores=0.59979904,
final_local_scores=0.7274614, pem=0.0, similarity=25.798613]
40: 5042481 Cuba
[pem_pos=0, final_scores=0.49934,
final_local_scores=0.48951477, pem=-0.23907764,

```
similarity=19.997952]
41: 2556569 Letitia_Vriesde
    [pem_pos=0, final_scores=0.70937526,
final_local_scores=0.8415183, pem=0.0, similarity=29.202297]
42: 26828 Suriname
    [pem_pos=0, final_scores=0.4443826,
final_local_scores=0.44434458, pem=0.0, similarity=17.349842]
43: 2736349 Inez_Turner
    [pem_pos=0, final_scores=0.55629784,
final_local_scores=0.73194104, pem=0.0, similarity=25.932293]
44: 15660 Jamaica
    [pem_pos=0, final_scores=0.46815526,
final_local_scores=0.46167326, pem=-0.33295843,
similarity=19.677643]
45: 11867 Germany
    [pem_pos=0, final_scores=0.79564184,
final_local_scores=0.79781836, pem=-0.15529114,
similarity=28.742702]
46: 9282173 Israel
    [pem_pos=0, final_scores=0.5212416,
final_local_scores=0.5405093, pem=-0.21279564,
similarity=21.376802]
47: 14532 Italy
    [pem_pos=0, final_scores=0.8032432,
final_local_scores=0.8044198, pem=-0.20612453,
similarity=29.216143]
48: 188171 Kenya
    [pem_pos=0, final_scores=0.59121424,
final_local_scores=0.6111072, pem=-0.27389407,
similarity=23.815847]
49: 188171 Kenya
    [pem_pos=0, final_scores=0.6087231,
final_local_scores=0.63023293, pem=-0.27389407,
similarity=24.386599]
50: 188171 Kenya
    [pem_pos=0, final_scores=0.6177419,
final_local_scores=0.6400845, pem=-0.27389407,
similarity=24.68059]
51: 25391 Russia
    [pem_pos=0, final_scores=0.7994421,
final_local_scores=0.8191589, pem=-0.20299491,
similarity=29.638966]
52: 3451457 Natalya_Shikolenko
    [pem_pos=0, final_scores=0.73940057,
final_local_scores=0.91367, pem=0.0, similarity=31.35545]
53: 3457 Belarus
    [pem_pos=0, final_scores=0.6309925,
final_local_scores=0.6321197, pem=-0.4220667,
similarity=25.24869]
54: 3910212 Silke_Renk
    [pem_pos=0, final_scores=0.7231198,
final_local_scores=0.8799649, pem=0.0, similarity=30.349619]
```

55: 11867 Germany
[pem_pos=0, final_scores=0.84151655,
final_local_scores=0.84792954, pem=-0.15529114,
similarity=30.238121]
56: 14532 Italy
[pem_pos=0, final_scores=0.7853916,
final_local_scores=0.7849196, pem=-0.20612453,
similarity=28.634218]
57: 7771507 Natalya_Torshina-Alimzhanova
[pem_pos=0, final_scores=0.48431933,
final_local_scores=0.66451454, pem=0.0, similarity=23.920153]
58: 16642 Kazakhstan
[pem_pos=0, final_scores=0.70542186,
final_local_scores=0.7164166, pem=-0.41320926,
similarity=27.71611]
59: 25391 Russia
[pem_pos=0, final_scores=0.79270905,
final_local_scores=0.81180406, pem=-0.20299491,
similarity=29.419483]
60: 14532 Italy
[pem_pos=0, final_scores=0.80344766,
final_local_scores=0.80464303, pem=-0.20612453,
similarity=29.222803]
61: 3434750 United_States
[pem_pos=0, final_scores=0.48163277,
final_local_scores=0.46407765, pem=-0.14075686,
similarity=18.704174]
62: 31717 United_Kingdom
[pem_pos=0, final_scores=0.57376724,
final_local_scores=0.570834, pem=-0.61114657,
similarity=24.448046]
63: 7754964 Luke_Kipkosgei
[pem_pos=0, final_scores=0.6371467,
final_local_scores=0.7987049, pem=0.0, similarity=27.924662]
64: 188171 Kenya
[pem_pos=0, final_scores=0.62349576,
final_local_scores=0.64636976, pem=-0.27389407,
similarity=24.868155]
65: 6815464 Alessandro_Lambruschini
[pem_pos=0, final_scores=0.6665942,
final_local_scores=0.80563766, pem=0.0, similarity=28.131548]
66: 14532 Italy
[pem_pos=0, final_scores=0.78684443,
final_local_scores=0.7865066, pem=-0.20612453,
similarity=28.681576]
67: 188171 Kenya
[pem_pos=0, final_scores=0.62486774,
final_local_scores=0.6478684, pem=-0.27389407,
similarity=24.912876]

Sample: b'638_GOLF'
GOLF - [1 1 BRITISH MASTERS 1] FIRST ROUND SCORES . [2 2 NOR

THAMPTON 2] , [3 3 England 3] 1996-08-28 Leading completed first-round scores in the rain-affected [4 4 British Masters 4] golf championship at [5 5 Collingtree Park 5] on Wednesday ([6 6 Britain 6] unless stated) : 66 Gavin Levenson ([7 7 S outh Africa 7]) 68 [8 8 Colin Montgomerie 8] 69 [9 9 Jose Co ceres 9] ([10 10 Argentina 10]) , [11 11 Raymond Russell 11] , [12 12 Robert Allenby 12] ([13 13 Australia 13]) , [14 14 David Gilford 14] , [15 15 Stuart Cage 15] , Mike Clayton ([1 6 16 Australia 16]) , [17 17 Mark Roe 17] , [18 18 Emanuele C anonica 18] ([19 19 Italy 19]) 70 Francisco Cea ([20 20 Sp ain 20]) , [21 21 David Howell 21] , [22 22 Peter Hedblom 22] ([23 23 Sweden 23]) 71 Steven Bottomley , [24 24 Ove Sellbe rg 24] ([25 25 Sweden 25]) , [26 26 Joakim Haeggman 26] ([2 7 27 Sweden 27]) , [28 28 Stephen Ames 28] (Trinidad and Tobago) , [29 29 Klas Eriksson 29] ([30 30 Sweden 30]) , [3 1 31 Roger Chapman 31] , [32 32 Mark Davis 32] , [33 33 Pierre Fulke 33] ([34 34 Sweden 34]) , Martin Gates , Anders Haglund ([35 35 Sweden 35]) 72 [36 36 Niclas Fasth 36] ([37 37 Sweden 37]) , [38 38 Michael Jonzon 38] ([39 39 Sweden 39]) , Chistian Cevaer ([40 40 France 40]) , [41 41 Thomas Bj orn 41] ([42 42 Denmark 42]) , [43 43 Tony Johnstone 43] ([44 44 Zimbabwe 44]) , [45 45 Padraig Harrington 45] ([46 46 Ireland 46]) , [47 47 Pedro Linhart 47] ([48 48 Spain 48]) , [49 49 David Carter 49] 73 [50 50 Ross McFarlane 50] , Domingo Hospital ([51 51 Spain 51]) , [52 52 Seve Bal lesteros 52] ([53 53 Spain 53]) , [54 54 Paul Broadhurst 54] , [55 55 Greg Turner 55] ([56 56 New Zealand 56]) , [57 57 M ike Harwood 57] ([58 58 Australia 58]) , [59 59 Brenden Pap pas 59] ([60 60 South Africa 60]) , Peter Teravainen ([61 6 1 U.S. 61]) , 1 Jean Van de Velde ([62 62 France 62]) , Oyvind Rojahn ([63 63 Norway 63]) , [64 64 Stephen McAllister 64] , Neal Briggs Note : 77 players to complete their first rounds on Thursday

- 1: 1929870 British_Masters
- 2: 63734 Northampton
- 3: 9316 England
- 4: 1929870 British_Masters
- 5: 11052655 Collingtree_Park
- 6: 31717 United_Kingdom
- 7: 17416221 South_Africa
- 8: 379616 Colin_Montgomerie
- 9: 2156041 José_Cóceres
- 10: 18951905 Argentina
- 11: 4375906 Raymond_Russell
- 12: 1983254 Robert_Allenby
- 13: 4689264 Australia
- 14: 2028738 David_Gilford
- 15: 5337735 Stuart_Cage
- 16: 4689264 Australia
- 17: 4353394 Mark_Roe
- 18: 2401223 Emanuele_Canonica
- 19: 14532 Italy

```
20: 26667 Spain
21: 1343807 David_Howell_(golfer)
22: 4377532 Peter_Hedblom
23: 5058739 Sweden
24: 5177483 Ove_Sellberg
25: 5058739 Sweden
26: 2032221 Joakim_Haeggman
27: 5058739 Sweden
28: 780871 Stephen_Ames
29: 28186295 Klas_Eriksson
30: 5058739 Sweden
31: 4380250 Roger_Chapman_(golfer)
32: 5120820 Mark_Davis_(golfer)
33: 2126111 Pierre_Fulke
34: 5058739 Sweden
35: 5058739 Sweden
36: 2033326 Niclas_Fasth
37: 5058739 Sweden
38: 5242970 Michael_Jonzon
39: 5058739 Sweden
40: 5843419 France
41: 1891067 Thomas_Bjørn
42: 76972 Denmark
43: 5217669 Tony_Johnstone
44: 34399 Zimbabwe
45: 1157761 Pádraig_Harrington
46: 14560 Republic_of_Ireland
47: 5285076 Pedro_Linhart
48: 26667 Spain
49: 4344819 David_Carter_(golfer)
50: 5190083 Ross_McFarlane
51: 26667 Spain
52: 740993 Seve_Ballesteros
53: 26667 Spain
54: 1688190 Paul_Broadhurst
55: 5276824 Greg_Turner
56: 4913064 New_Zealand
57: 5411795 Mike_Harwood
58: 4689264 Australia
59: 23282177 Brenden_Pappas
60: 17416221 South_Africa
61: 3434750 United_States
62: 5843419 France
63: 21241 Norway
64: 5266169 Stephen_McAllister
1: 2054489 Jean_van_de_Velde_(golfer)
    [winner: 1831915 Van_de_Velde, pem_pos=1,
final_scores=-0.3587745, final_local_scores=-0.2130005,
pem=-0.006211219, similarity=-2.2328699|
     RECALL MISS: 2054489 Jean_van_de_Velde_(golfer)]
1: 1929870 British_Masters
    [pem_pos=0, final_scores=0.40105253,
```

```
final_local_scores=0.47684366, pem=0.0, similarity=18.31968]
2: 63734 Northampton
    [pem_pos=0, final_scores=0.4331864,
final_local_scores=0.4866603, pem=-0.5110059,
similarity=21.391558]
3: 9316 England
    [pem_pos=0, final_scores=0.6000957,
final_local_scores=0.6228487, pem=-0.19053856,
similarity=23.712938]
4: 1929870 British_Masters
    [pem_pos=0, final_scores=0.45596433,
final_local_scores=0.53682643, pem=0.0, similarity=20.109684]
5: 11052655 Collingtree_Park
    [pem_pos=0, final_scores=0.35256708,
final_local_scores=0.49240643, pem=0.0, similarity=18.784103]
6: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.4823532,
final_local_scores=0.50179726, pem=-0.61114657,
similarity=22.387854]
7: 17416221 South_Africa
    [pem_pos=0, final_scores=0.53063506,
final_local_scores=0.5725923, pem=-0.26251382,
similarity=22.6046]
8: 379616 Colin_Montgomerie
    [pem_pos=0, final_scores=0.82085854,
final_local_scores=0.88073736, pem=0.0, similarity=30.372673]
9: 2156041 José_Cóceres
    [pem_pos=0, final_scores=0.70038337,
final_local_scores=0.77355653, pem=0.0, similarity=27.174183]
10: 18951905 Argentina
    [pem_pos=0, final_scores=0.6571327,
final_local_scores=0.69497097, pem=-0.30003807,
similarity=26.460686]
11: 4375906 Raymond_Russell
    [pem_pos=0, final_scores=0.7031614,
final_local_scores=0.778485, pem=0.0, similarity=27.321259]
12: 1983254 Robert_Allenby
    [pem_pos=0, final_scores=0.7564247,
final_local_scores=0.8426, pem=0.0, similarity=29.234577]
13: 4689264 Australia
    [pem_pos=0, final_scores=0.7161949,
final_local_scores=0.76001805, pem=-0.17168477,
similarity=27.703817]
14: 2028738 David_Gilford
    [pem_pos=0, final_scores=0.7952581,
final_local_scores=0.9027234, pem=0.0, similarity=31.028778]
15: 5337735 Stuart_Cage
    [pem_pos=0, final_scores=0.45585555,
final_local_scores=0.55887455, pem=0.0, similarity=20.767645]
16: 4689264 Australia
    [pem_pos=0, final_scores=0.6876178,
final_local_scores=0.72880185, pem=-0.17168477,
```

```
similarity=26.772263]
17: 4353394 Mark_Roe
    [pem_pos=0, final_scores=0.792458,
final_local_scores=0.8925248, pem=0.0, similarity=30.724432]
18: 2401223 Emanuele_Canonica
    [pem_pos=0, final_scores=0.6067663,
final_local_scores=0.729074, pem=0.0, similarity=25.846735]
19: 14532 Italy
    [pem_pos=0, final_scores=0.70336115,
final_local_scores=0.75731975, pem=-0.20612453,
similarity=27.810583]
20: 26667 Spain
    [pem_pos=0, final_scores=0.6119433,
final_local_scores=0.63211733, pem=-0.18857051,
similarity=23.97883]
21: 1343807 David_Howell_(golfer)
    [pem_pos=0, final_scores=0.65769964,
final_local_scores=0.72666574, pem=-0.78602827,
similarity=30.04941]
22: 4377532 Peter_Hedblom
    [pem_pos=0, final_scores=0.65880626,
final_local_scores=0.7338681, pem=0.0, similarity=25.989801]
23: 5058739 Sweden
    [pem_pos=0, final_scores=0.70641726,
final_local_scores=0.7210778, pem=-0.2135508,
similarity=26.769434]
24: 5177483 Ove_Sellberg
    [pem_pos=0, final_scores=0.63442504,
final_local_scores=0.7161631, pem=0.0, similarity=25.461449]
25: 5058739 Sweden
    [pem_pos=0, final_scores=0.7503404,
final_local_scores=0.76905715, pem=-0.2135508,
similarity=28.201235]
26: 2032221 Joakim_Haeggman
    [pem_pos=0, final_scores=0.66974,
final_local_scores=0.7433728, pem=0.0, similarity=26.27344]
27: 5058739 Sweden
    [pem_pos=0, final_scores=0.7255003,
final_local_scores=0.7419231, pem=-0.2135508,
similarity=27.3915]
28: 780871 Stephen_Ames
    [pem_pos=0, final_scores=0.63912755,
final_local_scores=0.7087222, pem=0.0, similarity=25.239397]
29: 28186295 Klas_Eriksson
    [pem_pos=0, final_scores=0.70864654,
final_local_scores=0.8216843, pem=0.0, similarity=28.61041]
30: 5058739 Sweden
    [pem_pos=0, final_scores=0.7121156,
final_local_scores=0.72730225, pem=-0.2135508,
similarity=26.955185]
31: 4380250 Roger_Chapman_(golfer)
    [pem_pos=1, final_scores=0.7114462,
```

```
final_local_scores=0.7857372, pem=-0.5573004,
similarity=30.568363]
32: 5120820 Mark_Davis_(golfer)
    [pem_pos=8, final_scores=0.49935973,
final_local_scores=0.57692647, pem=-2.4415264,
similarity=34.583748]
33: 2126111 Pierre_Fulke
    [pem_pos=0, final_scores=0.58726126,
final_local_scores=0.6601437, pem=0.0, similarity=23.789717]
34: 5058739 Sweden
    [pem_pos=0, final_scores=0.72453135,
final_local_scores=0.7408647, pem=-0.2135508,
similarity=27.359915]
35: 5058739 Sweden
    [pem_pos=0, final_scores=0.72590595,
final_local_scores=0.7423662, pem=-0.2135508,
similarity=27.404722]
36: 2033326 Niclas_Fasth
    [pem_pos=0, final_scores=0.80823046,
final_local_scores=0.89800876, pem=0.0, similarity=30.888086]
37: 5058739 Sweden
    [pem_pos=0, final_scores=0.7558513,
final_local_scores=0.775077, pem=-0.2135508,
similarity=28.380878]
38: 5242970 Michael_Jonzon
    [pem_pos=0, final_scores=0.6702006,
final_local_scores=0.7699604, pem=0.0, similarity=27.066868]
39: 5058739 Sweden
    [pem_pos=0, final_scores=0.7138979,
final_local_scores=0.72924924, pem=-0.2135508,
similarity=27.013287]
40: 5843419 France
    [pem_pos=0, final_scores=0.7621111,
final_local_scores=0.80606884, pem=-0.16240834,
similarity=29.027617]
41: 1891067 Thomas_Bjørn
    [pem_pos=0, final_scores=0.81392187,
final_local_scores=0.8892595, pem=0.0, similarity=30.626991]
42: 76972 Denmark
    [pem_pos=0, final_scores=0.56393903,
final_local_scores=0.57792634, pem=-0.26762357,
similarity=22.791567]
43: 5217669 Tony_Johnstone
    [pem_pos=0, final_scores=0.69572043,
final_local_scores=0.78614295, pem=0.0, similarity=27.549786]
44: 34399 Zimbabwe
    [pem_pos=0, final_scores=0.55149907,
final_local_scores=0.6147642, pem=-0.36737105,
similarity=24.433325]
45: 1157761 Pádraig_Harrington
    [pem_pos=0, final_scores=0.6511748,
final_local_scores=0.71178603, pem=0.0, similarity=25.330828]
```

```
46: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.5921779,
final_local_scores=0.6319304, pem=-1.5040284,
similarity=31.126919]
47: 5285076 Pedro_Linhart
    [pem_pos=0, final_scores=0.6218192,
final_local_scores=0.7122715, pem=0.0, similarity=25.345316]
48: 26667 Spain
    [pem_pos=0, final_scores=0.58565,
final_local_scores=0.60339594, pem=-0.18857051,
similarity=23.121727]
49: 4344819 David_Carter_(golfer)
    [pem_pos=4, final_scores=0.6102828,
final_local_scores=0.6671731, pem=-1.8641758,
similarity=34.137165]
50: 5190083 Ross_McFarlane
    [pem_pos=0, final_scores=0.6160016,
final_local_scores=0.6934403, pem=0.0, similarity=24.783356]
51: 26667 Spain
    [pem_pos=0, final_scores=0.60350424,
final_local_scores=0.62289894, pem=-0.18857051,
similarity=23.703735]
52: 740993 Seve_Ballesteros
    [pem_pos=0, final_scores=0.50230783,
final_local_scores=0.5305794, pem=0.0, similarity=19.92326]
53: 26667 Spain
    [pem_pos=0, final_scores=0.67510325,
final_local_scores=0.70111, pem=-0.18857051,
similarity=26.03771]
54: 1688190 Paul_Broadhurst
    [pem_pos=0, final_scores=0.84078914,
final_local_scores=0.9378906, pem=0.0, similarity=32.07824]
55: 5276824 Greg_Turner
    [pem_pos=0, final_scores=0.61737466,
final_local_scores=0.7002014, pem=0.0, similarity=24.985119]
56: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.49794537,
final_local_scores=0.531685, pem=-0.23795462,
similarity=21.250288]
57: 5411795 Mike_Harwood
    [pem_pos=0, final_scores=0.5365072,
final_local_scores=0.6068485, pem=0.0, similarity=22.199282]
58: 4689264 Australia
    [pem_pos=0, final_scores=0.6749315,
final_local_scores=0.71494406, pem=-0.17168477,
similarity=26.358719]
59: 23282177 Brenden_Pappas
    [pem_pos=0, final_scores=0.5339578,
final_local_scores=0.6366789, pem=0.0, similarity=23.089481]
60: 17416221 South_Africa
    [pem_pos=0, final_scores=0.6659551,
final_local_scores=0.7204089, pem=-0.26251382,
```

```
similarity=27.015743]
61: 3434750 United_States
    [pem_pos=0, final_scores=0.4484145,
final_local_scores=0.46760362, pem=-0.14075686,
similarity=18.809397]
62: 5843419 France
    [pem_pos=0, final_scores=0.6771168,
final_local_scores=0.7132254, pem=-0.16240834,
similarity=26.256983]
63: 21241 Norway
    [pem_pos=0, final_scores=0.77025706,
final_local_scores=0.82773495, pem=-0.22668563,
similarity=30.023726]
64: 5266169 Stephen_McAllister
    [pem_pos=0, final_scores=0.69672996,
final_local_scores=0.7849639, pem=0.0, similarity=27.5146]
```

Sample: b'639_CYCLING'

CYCLING - [1 1 KELLY 1] WINS WORLD TIME TRIAL TITLE . [2 2 MA
NCHESTER 2] , [3 3 England 3] 1996-08-28 [4 4 Shane Kelly 4]
of [5 5 Australia 5] retained the world one-kilometre time
trial title at the world track championships on Wednesday ,
with a track record time of one minute 2.777 seconds . [6 6 Ke
lly 6] averaged 57.345 kph to beat Soren Lausberg of [7 7 Ge
rmany 7] by eighteen hundredths of a second . The bronze
medal went to another [8 8 German 8] , [9 9 Jan Van Eijden 9]
, in 1:04.541 .

```
1: 4167427 Shane_Kelly
2: 20206 Manchester
3: 9316 England
4: 4167427 Shane_Kelly
5: 4689264 Australia
6: 4167427 Shane_Kelly
7: 11867 Germany
8: 11867 Germany
9: 9145159 Jan_van_Eijden
1: 4167427 Shane_Kelly
    [pem_pos=0, final_scores=0.52460295,
final_local_scores=0.58506095, pem=0.0, similarity=21.549099]
2: 20206 Manchester
    [pem_pos=0, final_scores=0.54277354,
final_local_scores=0.566048, pem=-0.3209542,
similarity=22.727114]
3: 9316 England
    [pem_pos=0, final_scores=0.63266367,
final_local_scores=0.67214274, pem=-0.19053856,
similarity=25.183971]
4: 4167427 Shane_Kelly
    [pem_pos=0, final_scores=0.55580205,
final_local_scores=0.6191412, pem=0.0, similarity=22.566122]
5: 4689264 Australia
    [pem_pos=0, final_scores=0.7413642,
```

```
final_local_scores=0.79957724, pem=-0.17168477,
similarity=28.884342]
6: 4167427 Shane_Kelly
    [pem_pos=0, final_scores=0.950314,
final_local_scores=1.0500855, pem=0.0, similarity=35.42636]
7: 11867 Germany
    [pem_pos=0, final_scores=0.86285836,
final_local_scores=0.92050654, pem=-0.15529114,
similarity=32.40396]
8: 11867 Germany
    [pem_pos=1, final_scores=0.744599,
final_local_scores=0.7913261, pem=-1.0465823,
similarity=33.39594]
9: 9145159 Jan_van_Eijden
    [pem_pos=0, final_scores=0.57200587,
final_local_scores=0.71902645, pem=0.0, similarity=25.546896]
```

Sample: b'640_CYCLING'

CYCLING - [1 1 VAN HEESWIJK 1] WINS SECOND STAGE OF TOUR OF NETHERLANDS . [2 2 ALMERE 2] , [3 3 Netherlands 3] 1996-08-28 Leading placings in the 195-km second stage of the [4 4 Tour of the Netherlands 4] between [5 5 Haarlem 5] and [6 6 Almere 6] on Wednesday : 1. [7 7 Max van Heeswijk 7] ([8 8 Netherlands 8]) [9 9 Motorola 9] 4 hours 39 minutes six seconds . 2. [10 10 Johan Capiot 10] ([11 11 Belgium 11]) [12 12 Collstrop 12] 3. Sven Teutenberg ([13 13 Germany 13]) 1 U.S. Postal 4. [14 14 Erik Zabel 14] ([15 15 Germany 15]) Telekom 5. Federico Colonna ([16 16 Italy 16]) Mapei 6. [17 17 Jans Koerts 17] ([18 18 Netherlands 18]) Palmans 7. Michel Zanoli ([19 19 Netherlands 19]) MX Onda 8 .Giuseppe Citterio ([20 20 Italy 20]) Aki 9 .Robbie McEwen ([21 21 Australia 21]) [22 22 Rabobank 22] 10. Kaspars Ozers ([23 23 Latvia 23]) [24 24 Motorola 24] all same time Leading overall placings after two stages : 1. Colonna 8:22:00 2. Van Heeswijk 1 second behind 3. [25 25 McEwen 25] same time 4. Teutenberg 4 seconds 5. [26 26 Capiot 26] 5 6. [27 27 Koerts 27] 7 7. Ozers 8 8. Gianluca Corini ([28 28 Italy 28]) Aki same time 9. [29 29 Lance Armstrong 29] ([30 30 U.S. 30]) [31 31 Motorola 31] 9 10. [32 32 George Hincapie 32] ([33 33 U.S. 33]) [34 34 Motorola 34] same time

```
1: 6558171 Max_van_Heeswijk
2: 118444 Almere
3: 21148 Netherlands
4: 11065220 Ronde_van_Nederland
5: 63278 Haarlem
6: 118444 Almere
7: 6558171 Max_van_Heeswijk
8: 21148 Netherlands
9: 2290911 7-Eleven_(cycling_team)
10: 13573637 Johan_Capiot
11: 3343 Belgium
12: 6932095 Cycle_Collstrop
```

13: 11867 Germany
14: 921007 Erik_Zabel
15: 11867 Germany
16: 14532 Italy
17: 17022543 Jans_Koerts
18: 21148 Netherlands
19: 21148 Netherlands
20: 14532 Italy
21: 4689264 Australia
22: 2354465 Belkin_Pro_Cycling_Team
23: 17514 Latvia
24: 2290911 7-Eleven_(cycling_team)
25: 1484348 Robbie_McEwen
26: 13573637 Johan_Capiot
27: 17022543 Jans_Koerts
28: 14532 Italy
29: 23243880 Lance_Armstrong
30: 3434750 United_States
31: 2290911 7-Eleven_(cycling_team)
32: 1257631 George_Hincapie
33: 3434750 United_States
34: 2290911 7-Eleven_(cycling_team)
1: 796027 U.S._Postal_Service_Pro_Cycling_Team
 [winner: 1148045 Postal_(video_game), pem_pos=0,
final_scores=-1.017054, final_local_scores=-0.97273237,
pem=-0.8678312, similarity=-20.21916]
 RECALL MISS: 796027
U.S._Postal_Service_Pro_Cycling_Team]
 [winner: 3268207 United_States_Reports, pem_pos=1,
final_scores=-1.2473245, final_local_scores=-1.1969123,
pem=-2.4891596, similarity=-18.092102]
 RECALL MISS: 796027 U.S._Postal_Service_Pro_Cycling_Team]
1: 6558171 Max_van_Heeswijk
 [pem_pos=0, final_scores=0.4411958,
final_local_scores=0.5274328, pem=0.0, similarity=19.82936]
2: 118444 Almere
 [pem_pos=0, final_scores=0.40330207,
final_local_scores=0.42916912, pem=0.0, similarity=16.896975]
3: 21148 Netherlands
 [pem_pos=2, final_scores=0.5849722,
final_local_scores=0.5694499, pem=-0.24940795,
similarity=22.439554]
4: 11065220 Ronde_van_Nederland
 [pem_pos=0, final_scores=0.52899015,
final_local_scores=0.6174097, pem=0.0, similarity=22.51445]
5: 63278 Haarlem
 [pem_pos=0, final_scores=0.42090434,
final_local_scores=0.45557588, pem=0.0, similarity=17.685005]
6: 118444 Almere
 [pem_pos=0, final_scores=0.41783887,
final_local_scores=0.44504833, pem=0.0, similarity=17.370842]
7: 6558171 Max_van_Heeswijk

```
[pem_pos=0, final_scores=0.5215356,
final_local_scores=0.6151918, pem=0.0, similarity=22.448263]
8: 21148 Netherlands
    [pem_pos=2, final_scores=0.637759,
final_local_scores=0.6271114, pem=-0.24940795,
similarity=24.160286]
9: 2290911 7-Eleven_(cycling_team)
    [pem_pos=1, final_scores=0.7428626,
final_local_scores=0.82460654, pem=-1.7647687,
similarity=38.294704]
10: 13573637 Johan_Capiot
    [pem_pos=0, final_scores=0.67240834,
final_local_scores=0.8225959, pem=0.0, similarity=28.637615]
11: 3343 Belgium
    [pem_pos=0, final_scores=0.73987883,
final_local_scores=0.7395712, pem=-0.24180195,
similarity=27.47495]
12: 6932095 Cycle_Collstrop
    [pem_pos=0, final_scores=0.5561804,
final_local_scores=0.6735936, pem=0.0, similarity=24.19109]
13: 11867 Germany
    [pem_pos=0, final_scores=0.83828,
final_local_scores=0.869389, pem=-0.15529114,
similarity=30.878511]
14: 921007 Erik_Zabel
    [pem_pos=0, final_scores=0.8068897,
final_local_scores=0.8840721, pem=0.0, similarity=30.472187]
15: 11867 Germany
    [pem_pos=0, final_scores=0.88017374,
final_local_scores=0.91515154, pem=-0.15529114,
similarity=32.24416]
16: 14532 Italy
    [pem_pos=0, final_scores=0.73808414,
final_local_scores=0.78062886, pem=-0.20612453,
similarity=28.506172]
17: 17022543 Jans_Koerts
    [pem_pos=0, final_scores=0.61833465,
final_local_scores=0.73705226, pem=0.0, similarity=26.084822]
18: 21148 Netherlands
    [pem_pos=2, final_scores=0.65683544,
final_local_scores=0.6479495, pem=-0.24940795,
similarity=24.782139]
19: 21148 Netherlands
    [pem_pos=2, final_scores=0.6700829,
final_local_scores=0.6624204, pem=-0.24940795,
similarity=25.213978]
20: 14532 Italy
    [pem_pos=0, final_scores=0.7498931,
final_local_scores=0.79352844, pem=-0.20612453,
similarity=28.891123]
21: 4689264 Australia
    [pem_pos=0, final_scores=0.64034957,
```

```
final_local_scores=0.7080671, pem=-0.17168477,  
similarity=26.153498]  
22: 2354465 Belkin_Pro_Cycling_Team  
    [pem_pos=0, final_scores=0.64499074,  
final_local_scores=0.7097797, pem=0.0, similarity=25.270954]  
23: 17514 Latvia  
    [pem_pos=0, final_scores=0.5992076,  
final_local_scores=0.6414314, pem=-0.35144013,  
similarity=25.142492]  
24: 2290911 7-Eleven_(cycling_team)  
    [pem_pos=1, final_scores=0.76579744,  
final_local_scores=0.84965944, pem=-1.7647687,  
similarity=39.04233]  
25: 1484348 Robbie_McEwen  
    [pem_pos=7, final_scores=0.7147003,  
final_local_scores=0.8033509, pem=-3.0529037,  
similarity=44.66547]  
26: 13573637 Johan_Capiot  
    [pem_pos=0, final_scores=0.6222181,  
final_local_scores=0.76777065, pem=0.0, similarity=27.00152]  
27: 17022543 Jans_Koerts  
    [pem_pos=0, final_scores=0.49600214,  
final_local_scores=0.6034226, pem=0.0, similarity=22.097046]  
28: 14532 Italy  
    [pem_pos=0, final_scores=0.72975606,  
final_local_scores=0.77153164, pem=-0.20612453,  
similarity=28.234694]  
29: 23243880 Lance_Armstrong  
    [pem_pos=0, final_scores=0.5750061,  
final_local_scores=0.6292862, pem=0.0, similarity=22.86887]  
30: 3434750 United_States  
    [pem_pos=0, final_scores=0.46596223,  
final_local_scores=0.49381346, pem=-0.14075686,  
similarity=19.591549]  
31: 2290911 7-Eleven_(cycling_team)  
    [pem_pos=1, final_scores=0.742317,  
final_local_scores=0.8240106, pem=-1.7647687,  
similarity=38.27692]  
32: 1257631 George_Hincapie  
    [pem_pos=0, final_scores=0.7561448,  
final_local_scores=0.8330079, pem=0.0, similarity=28.94833]  
33: 3434750 United_States  
    [pem_pos=0, final_scores=0.5086392,  
final_local_scores=0.54043156, pem=-0.14075686,  
similarity=20.982727]  
34: 2290911 7-Eleven_(cycling_team)  
    [pem_pos=1, final_scores=0.65192443,  
final_local_scores=0.7252704, pem=-1.7647687,  
similarity=35.330315]
```

Sample: b'641_CYCLING'

CYCLING - [1 1 VAN HEESWIJK 1] WINS TOUR OF NETHERLANDS

SECOND STAGE . [2 2 ALMERE 2] , [3 3 Netherlands 3]

1996-08-28 Leading results and overall standings after the 195 kilometre second stage of the [4 4 Tour of the Netherlands 4] between [5 5 Haarlem 5] and [6 6 Almere 6] on Wednesday .

1. [7 7 Max van Heeswijk 7] ([8 8 Netherlands 8]) [9 9 Moto rola 9] 4 hours 39 minutes six seconds . 2. [10 10 Johan Capi ot 10] ([11 11 Belgium 11]) [12 12 Collstrop 12] 3. Sven Teutenberg ([13 13 Germany 13]) 1 U.S. Postal 4. [14 14 Erik Zabel 14] ([15 15 Germany 15]) Telekom 5. Federico Colonna ([16 16 Italy 16]) Mapei 6. [17 17 Jans Koerts 17] ([18 18 Netherlands 18]) Palmans 7. Michel Zanoli ([19 19 Netherlan ds 19]) MX Onda 8. Giuseppe Citterio ([20 20 Italy 20]) Aki 9. [21 21 Robbie McEwen 21] ([22 22 Australia 22]) [23 23 Rabobank 23] 10. Kaspars Ozers ([24 24 Latvia 24]) [25 25 Motorola 25] all same time . Leading overall standings after second stage . 1. Colonna 8:22:00 2. Van Heeswijk 1 second behind 3. [26 26 McEwen 26] same time 4. Teutenberg 4 seconds 5. [27 27 Capiot 27] 5 6. [28 28 Koerts 28] 7 7. Ozers 8 8. Gianluca Corini ([29 29 Italy 29]) Aki same time 9. [30 30 Lance Armstrong 30] ([31 31 U.S. 31]) [32 32 Motorola 32] 9 10. [33 33 George Hincapie 33] ([34 34 U.S. 34]) [35 35 Mot orola 35] same time .

1: 6558171 Max_van_Heeswijk

2: 118444 Almere

3: 21148 Netherlands

4: 11065220 Ronde_van_Nederland

5: 63278 Haarlem

6: 118444 Almere

7: 6558171 Max_van_Heeswijk

8: 21148 Netherlands

9: 2290911 7-Eleven_(cycling_team)

10: 13573637 Johan_Capiot

11: 3343 Belgium

12: 6932095 Cycle_Collstrop

13: 11867 Germany

14: 921007 Erik_Zabel

15: 11867 Germany

16: 14532 Italy

17: 17022543 Jans_Koerts

18: 21148 Netherlands

19: 21148 Netherlands

20: 14532 Italy

21: 1484348 Robbie_McEwen

22: 4689264 Australia

23: 2354465 Belkin_Pro_Cycling_Team

24: 17514 Latvia

25: 2290911 7-Eleven_(cycling_team)

26: 1484348 Robbie_McEwen

27: 13573637 Johan_Capiot

28: 17022543 Jans_Koerts

29: 14532 Italy

30: 23243880 Lance_Armstrong

31: 3434750 United_States
32: 2290911 7-Eleven_(cycling_team)
33: 1257631 George_Hincapie
34: 3434750 United_States
35: 2290911 7-Eleven_(cycling_team)
1: 796027 U.S._Postal_Service_Pro_Cycling_Team
[winner: 1148045 Postal_(video_game), pem_pos=0,
final_scores=-1.017434, final_local_scores=-0.97244424,
pem=-0.8678312, similarity=-20.210564]
RECALL MISS: 796027
U.S._Postal_Service_Pro_Cycling_Team]
[winner: 3268207 United_States_Reports, pem_pos=1,
final_scores=-1.2491195, final_local_scores=-1.1968162,
pem=-2.4891596, similarity=-18.089233]
RECALL MISS: 796027 U.S._Postal_Service_Pro_Cycling_Team]
1: 6558171 Max_van_Heeswijk
[pem_pos=0, final_scores=0.43897372,
final_local_scores=0.523619, pem=0.0, similarity=19.715548]
2: 118444 Almere
[pem_pos=0, final_scores=0.39739728,
final_local_scores=0.42583752, pem=0.0, similarity=16.797554]
3: 21148 Netherlands
[pem_pos=2, final_scores=0.56911844,
final_local_scores=0.5584848, pem=-0.24940795,
similarity=22.112333]
4: 11065220 Ronde_van_Nederland
[pem_pos=0, final_scores=0.60047704,
final_local_scores=0.69062966, pem=0.0, similarity=24.69948]
5: 63278 Haarlem
[pem_pos=0, final_scores=0.49410105,
final_local_scores=0.5393477, pem=0.0, similarity=20.184925]
6: 118444 Almere
[pem_pos=0, final_scores=0.45431447,
final_local_scores=0.48801088, pem=0.0, similarity=18.652931]
7: 6558171 Max_van_Heeswijk
[pem_pos=0, final_scores=0.5518551,
final_local_scores=0.64692473, pem=0.0, similarity=23.395237]
8: 21148 Netherlands
[pem_pos=2, final_scores=0.6438483,
final_local_scores=0.6401158, pem=-0.24940795,
similarity=24.548365]
9: 2290911 7-Eleven_(cycling_team)
[pem_pos=1, final_scores=0.77588814,
final_local_scores=0.85713017, pem=-1.7647687,
similarity=39.265274]
10: 13573637 Johan_Capiot
[pem_pos=0, final_scores=0.67194486,
final_local_scores=0.81802475, pem=0.0, similarity=28.501204]
11: 3343 Belgium
[pem_pos=0, final_scores=0.7375352,
final_local_scores=0.74223304, pem=-0.24180195,
similarity=27.554384]

12: 6932095 Cycle_Collstrop
[pem_pos=0, final_scores=0.5671973,
final_local_scores=0.6796962, pem=0.0, similarity=24.373203]
13: 11867 Germany
[pem_pos=0, final_scores=0.8314356,
final_local_scores=0.86867064, pem=-0.15529114,
similarity=30.857075]
14: 921007 Erik_Zabel
[pem_pos=0, final_scores=0.82189924,
final_local_scores=0.89026564, pem=0.0, similarity=30.657015]
15: 11867 Germany
[pem_pos=0, final_scores=0.87224954,
final_local_scores=0.9132536, pem=-0.15529114,
similarity=32.18752]
16: 14532 Italy
[pem_pos=0, final_scores=0.731512,
final_local_scores=0.77789116, pem=-0.20612453,
similarity=28.424475]
17: 17022543 Jans_Koerts
[pem_pos=0, final_scores=0.6224586,
final_local_scores=0.74111676, pem=0.0, similarity=26.206116]
18: 21148 Netherlands
[pem_pos=2, final_scores=0.6516955,
final_local_scores=0.64868766, pem=-0.24940795,
similarity=24.804167]
19: 21148 Netherlands
[pem_pos=2, final_scores=0.6541592,
final_local_scores=0.65137887, pem=-0.24940795,
similarity=24.884478]
20: 14532 Italy
[pem_pos=0, final_scores=0.7665716,
final_local_scores=0.81618845, pem=-0.20612453,
similarity=29.567343]
21: 1484348 Robbie_McEwen
[pem_pos=0, final_scores=0.7309306,
final_local_scores=0.8016145, pem=0.0, similarity=28.01149]
22: 4689264 Australia
[pem_pos=0, final_scores=0.6978076,
final_local_scores=0.77584267, pem=-0.17168477,
similarity=28.176054]
23: 2354465 Belkin_Pro_Cycling_Team
[pem_pos=0, final_scores=0.6253001,
final_local_scores=0.68239903, pem=0.0, similarity=24.453861]
24: 17514 Latvia
[pem_pos=0, final_scores=0.61984354,
final_local_scores=0.6684287, pem=-0.35144013,
similarity=25.948145]
25: 2290911 7-Eleven_(cycling_team)
[pem_pos=1, final_scores=0.7650288,
final_local_scores=0.845268, pem=-1.7647687,
similarity=38.91128]
26: 1484348 Robbie_McEwen

```
[pem_pos=0, final_scores=1.2567749,
final_local_scores=1.3853555, pem=0.0, similarity=45.43148]
27: 13573637 Johan_Capiot
    [pem_pos=0, final_scores=0.64060223,
final_local_scores=0.7837877, pem=0.0, similarity=27.479502]
28: 17022543 Jans_Koerts
    [pem_pos=0, final_scores=0.5041431,
final_local_scores=0.6118751, pem=0.0, similarity=22.349287]
29: 14532 Italy
    [pem_pos=0, final_scores=0.7202448,
final_local_scores=0.7655836, pem=-0.20612453,
similarity=28.057192]
30: 23243880 Lance_Armstrong
    [pem_pos=0, final_scores=0.5831358,
final_local_scores=0.6327177, pem=0.0, similarity=22.971272]
31: 3434750 United_States
    [pem_pos=0, final_scores=0.46606493,
final_local_scores=0.5001064, pem=-0.14075686,
similarity=19.779345]
32: 2290911 7-Eleven_(cycling_team)
    [pem_pos=1, final_scores=0.7191204,
final_local_scores=0.79512, pem=-1.7647687,
similarity=37.414764]
33: 1257631 George_Hincapie
    [pem_pos=0, final_scores=0.74218243,
final_local_scores=0.80911857, pem=0.0, similarity=28.235424]
34: 3434750 United_States
    [pem_pos=0, final_scores=0.5071035,
final_local_scores=0.5449348, pem=-0.14075686,
similarity=21.117113]
35: 2290911 7-Eleven_(cycling_team)
    [pem_pos=1, final_scores=0.66838396,
final_local_scores=0.7396982, pem=-1.7647687,
similarity=35.760868]
```

Sample: b'642_CYCLING'

CYCLING - WORLD CHAMPIONSHIPS RESULTS . [1 1 MANCHESTER 1] ,
[2 2 England 2] 1996-08-28 Results from the world track
cycling championships on Wednesday : Men 's individual
pursuit : Selected result from first round [3 3 Chris Boardman
3] ([4 4 Britain 4]) 4:13.353 seconds (world record)
caught Jens Lehman ([5 5 Germany 5]) Quarter-finals [6 6 Bo
ardman 6] 4:14.784 caught Edouard Gritson ([7 7 Russia 7]) [
8 8 Francis Moreau 8] ([9 9 France 9]) 4:16.274 beat Heiko
Szonn ([10 10 Germany 10]) 4:21.715 [11 11 Andrea Collinelli
11] ([12 12 Italy 12]) 4:17.551 beat Michael Sandstod ([13
13 Denmark 13]) 4:24.660 [14 14 Alexei Markov 14] ([15 15 Ru
ssia 15]) 4:19.762 beat Mariano Friedick ([16 16 U.S. 16])
4:20.241 one kilometre time-trial final 1. [17 17 Shane Kelly
17] ([18 18 Australia 18]) 1 minute 02.777 seconds 2. [19 19
Soren Lausberg 19] ([20 20 Germany 20]) 1:02.795 3. Jan Van
Eiden ([21 21 Germany 21]) 1:04.541 4. Herve Thuet ([22 22

France 22]) 1:04.732 5. Grzegorz Krejner ([23 23 Poland 23]) 1:04.834 6. Ainars Kiksis ([24 24 Latvia 24]) 1:04.896 7. Dimitrios Georgalis ([25 25 Greece 25]) 1:05.022 8. Jose Moreno ([26 26 Spain 26]) 1:05.219 9. Keiji Kojima ([27 27 Japan 27]) 1:05.300 10. Graham Sharman ([28 28 Australia 28]) 1:05.406 11. [29 29 Jose Escuredo 29] ([30 30 Spain 30]) 1:05.731 12. [31 31 Craig MacLean 31] ([32 32 Britain 32]) 1:05.735 13. Christian Meidlinger ([33 33 Austria 33]) 1:05.850 14. Darren McKenzie-Potter ([34 34 New Zealand 34]) 1:06.289 15. Masanaga Shiohara ([35 35 Japan 35]) 1:06.615 16. Jean-Pierre Van Zyl ([36 36 South Africa 36]) 1:07.258 [37 37 Keirin 37] final 1. [38 38 Marty Nothstein 38] ([39 39 U.S. 39]) , last 200 metres in 10.982 seconds . 2. [40 40 Gary Neiwand 40] ([41 41 Australia 41]) 3. [42 42 Frederic Magne 42] ([43 43 France 43]) 4. [44 44 Pavel Buran 44] ([45 45 Czech Republic 45]) 5. [46 46 Michael Hubner 46] ([47 47 Germany 47]) 6. [48 48 Laurent Gane 48] ([49 49 France 49]) Madison final (50 kms) 1. Silvio Martinelli - [50 50 Marco Villa 50] ([51 51 Italy 51]) 34 points , in a time of 55 minutes 47.4 seconds 2. [52 52 Scott McGrory 52] - [53 53 Stephen Pate 53] ([54 54 Australia 54]) 25 3. [55 55 Andreas Kappes 55] - [56 56 Carsten Wolf 56] ([57 57 Germany 57]) 23 4. [58 58 Kurt Betschart 58] - [59 59 Bruno Risi 59] ([60 60 Switzerland 60]) 22 5. [61 61 Gabriel Curuchet 61] - [62 62 Juan Curuchet 62] ([63 63 Argentine 63]) 15 6. Peter Pieters - Tomas Post ([64 64 Netherlands 64]) 14 7. Jimmie Madsen - Jens Veggerby ([65 65 Denmark 65]) 14 8. Isaac Galvez-Lopez - [66 66 Juan Llaneras 66] ([67 67 Spain 67]) 11 9. Wolfgang Kotzmann - Franz Stocher ([68 68 Austria 68]) 5 10. Christophe Capelle - Jean-Michel Monin ([69 69 France 69]) 2
1: 20206 Manchester
2: 9316 England
3: 917770 Chris_Boardman
4: 31717 United_Kingdom
5: 11867 Germany
6: 917770 Chris_Boardman
7: 25391 Russia
8: 18894901 Francis_Moreau
9: 5843419 France
10: 11867 Germany
11: 17826490 Andrea_Collinelli
12: 14532 Italy
13: 76972 Denmark
14: 12752327 Alexei_Markov
15: 25391 Russia
16: 3434750 United_States
17: 4167427 Shane_Kelly
18: 4689264 Australia
19: 6691361 Soren_Lausberg
20: 11867 Germany
21: 11867 Germany

22: 5843419 France
23: 22936 Poland
24: 17514 Latvia
25: 12108 Greece
26: 26667 Spain
27: 15573 Japan
28: 4689264 Australia
29: 14689876 José_Antonio_Escuredo
30: 26667 Spain
31: 6376819 Craig_MacLean
32: 31717 United_Kingdom
33: 26964606 Austria
34: 4913064 New_Zealand
35: 15573 Japan
36: 17416221 South_Africa
37: 30619 Keirin
38: 5940960 Marty_Nothstein
39: 3434750 United_States
40: 11438632 Gary_Neiwand
41: 4689264 Australia
42: 23236017 Frédéric_Magné
43: 5843419 France
44: 16745041 Pavel_Buran
45: 5321 Czech_Republic
46: 19109147 Michael_Hübner
47: 11867 Germany
48: 16732280 Laurent_Gané
49: 5843419 France
50: 11438203 Marco_Villa
51: 14532 Italy
52: 22665655 Scott_McGrory
53: 23728468 Stephen_Pate
54: 4689264 Australia
55: 11003628 Andreas_Kappes
56: 14671526 Carsten_Wolf
57: 11867 Germany
58: 20417509 Kurt_Betschart
59: 15169629 Bruno_Risi
60: 26748 Switzerland
61: 11435812 Gabriel_Curuchet
62: 11435903 Juan_Curuchet
63: 18951905 Argentina
64: 21148 Netherlands
65: 76972 Denmark
66: 12935832 Joan_Llaneras
67: 26667 Spain
68: 26964606 Austria
69: 5843419 France
1: 20206 Manchester
[pem_pos=0, final_scores=0.5568971,
final_local_scores=0.57267827, pem=-0.3209542,
similarity=22.924974]

2: 9316 England
[pem_pos=0, final_scores=0.6792086,
final_local_scores=0.6958717, pem=-0.19053856,
similarity=25.892092]
3: 917770 Chris_Boardman
[pem_pos=0, final_scores=0.7157325,
final_local_scores=0.81412697, pem=0.0, similarity=28.384886]
4: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5792067,
final_local_scores=0.57075065, pem=-0.61114657,
similarity=24.44556]
5: 11867 Germany
[pem_pos=0, final_scores=0.7926193,
final_local_scores=0.78375113, pem=-0.15529114,
similarity=28.322908]
6: 917770 Chris_Boardman
[pem_pos=0, final_scores=1.2821327,
final_local_scores=1.4381881, pem=0.0, similarity=47.008114]
7: 25391 Russia
[pem_pos=0, final_scores=0.7838282,
final_local_scores=0.81495315, pem=-0.20299491,
similarity=29.513458]
8: 18894901 Francis_Moreau
[pem_pos=0, final_scores=0.7148983,
final_local_scores=0.85800487, pem=0.0, similarity=29.69429]
9: 5843419 France
[pem_pos=0, final_scores=0.7815879,
final_local_scores=0.7773642, pem=-0.16240834,
similarity=28.171013]
10: 11867 Germany
[pem_pos=0, final_scores=0.91955596,
final_local_scores=0.9224101, pem=-0.15529114,
similarity=32.460766]
11: 17826490 Andrea_Collinelli
[pem_pos=0, final_scores=0.5503402,
final_local_scores=0.6832459, pem=0.0, similarity=24.479134]
12: 14532 Italy
[pem_pos=0, final_scores=0.77971464,
final_local_scores=0.78973466, pem=-0.20612453,
similarity=28.777908]
13: 76972 Denmark
[pem_pos=0, final_scores=0.5920442,
final_local_scores=0.5892777, pem=-0.26762357,
similarity=23.130312]
14: 12752327 Alexei_Markov
[pem_pos=0, final_scores=0.5030216,
final_local_scores=0.6587951, pem=0.0, similarity=23.749474]
15: 25391 Russia
[pem_pos=0, final_scores=0.83227307,
final_local_scores=0.8678718, pem=-0.20299491,
similarity=31.092655]
16: 3434750 United_States

```
[pem_pos=0, final_scores=0.5421832,
final_local_scores=0.529216, pem=-0.14075686,
similarity=20.648033]
17: 4167427 Shane_Kelly
    [pem_pos=0, final_scores=0.66247797,
final_local_scores=0.8033068, pem=0.0, similarity=28.06199]
18: 4689264 Australia
    [pem_pos=0, final_scores=0.67845666,
final_local_scores=0.6849123, pem=-0.17168477,
similarity=25.462511]
19: 6691361 Sören_Lausberg
    [pem_pos=0, final_scores=0.69953495,
final_local_scores=0.8842264, pem=0.0, similarity=30.476793]
20: 11867 Germany
    [pem_pos=0, final_scores=0.79704934,
final_local_scores=0.7885902, pem=-0.15529114,
similarity=28.467316]
21: 11867 Germany
    [pem_pos=0, final_scores=0.85301405,
final_local_scores=0.84972316, pem=-0.15529114,
similarity=30.291643]
22: 5843419 France
    [pem_pos=0, final_scores=0.74599916,
final_local_scores=0.7384889, pem=-0.16240834,
similarity=27.010897]
23: 22936 Poland
    [pem_pos=0, final_scores=0.60595006,
final_local_scores=0.59491354, pem=-0.15333635,
similarity=22.676987]
24: 17514 Latvia
    [pem_pos=0, final_scores=0.6359363,
final_local_scores=0.6412829, pem=-0.35144013,
similarity=25.138062]
25: 12108 Greece
    [pem_pos=0, final_scores=0.8228151,
final_local_scores=0.84744537, pem=-0.27607638,
similarity=30.88052]
26: 26667 Spain
    [pem_pos=0, final_scores=0.6570689,
final_local_scores=0.65041536, pem=-0.18857051,
similarity=24.52488]
27: 15573 Japan
    [pem_pos=0, final_scores=0.70605594,
final_local_scores=0.7356332, pem=-0.13549899,
similarity=26.779339]
28: 4689264 Australia
    [pem_pos=0, final_scores=0.7577072,
final_local_scores=0.7714815, pem=-0.17168477,
similarity=28.04591]
29: 14689876 José_Antonio_Escuredo
    [pem_pos=0, final_scores=0.548329,
final_local_scores=0.66983724, pem=0.0, similarity=24.078993]
```

30: 26667 Spain
[pem_pos=0, final_scores=0.5863612,
final_local_scores=0.57317793, pem=-0.18857051,
similarity=22.219961]
31: 6376819 Craig_MacLean
[pem_pos=0, final_scores=0.6320365,
final_local_scores=0.74359566, pem=0.0, similarity=26.28009]
32: 31717 United_Kingdom
[pem_pos=0, final_scores=0.56954956,
final_local_scores=0.5602017, pem=-0.61114657,
similarity=24.130758]
33: 26964606 Austria
[pem_pos=0, final_scores=0.6659973,
final_local_scores=0.6474568, pem=-0.25262418,
similarity=24.784925]
34: 4913064 New_Zealand
[pem_pos=0, final_scores=0.59442455,
final_local_scores=0.6111499, pem=-0.23795462,
similarity=23.62168]
35: 15573 Japan
[pem_pos=0, final_scores=0.7774604,
final_local_scores=0.8136318, pem=-0.13549899,
similarity=29.106972]
36: 17416221 South_Africa
[pem_pos=0, final_scores=0.60374737,
final_local_scores=0.6339871, pem=-0.26251382,
similarity=24.436745]
37: 30619 Keirin
[pem_pos=0, final_scores=0.27544862,
final_local_scores=0.35295677, pem=0.0, similarity=14.622647]
38: 5940960 Marty_Nothstein
[pem_pos=0, final_scores=0.7216124,
final_local_scores=0.8518212, pem=0.0, similarity=29.509756]
39: 3434750 United_States
[pem_pos=0, final_scores=0.52605546,
final_local_scores=0.5115989, pem=-0.14075686,
similarity=20.122301]
40: 11438632 Gary_Neiwand
[pem_pos=0, final_scores=0.6836965,
final_local_scores=0.82810235, pem=0.0, similarity=28.801939]
41: 4689264 Australia
[pem_pos=0, final_scores=0.8332222,
final_local_scores=0.85397017, pem=-0.17168477,
similarity=30.507536]
42: 23236017 Frédéric_Magné
[pem_pos=0, final_scores=0.6785137,
final_local_scores=0.8803753, pem=0.0, similarity=30.361868]
43: 5843419 France
[pem_pos=0, final_scores=0.74552876,
final_local_scores=0.737975, pem=-0.16240834,
similarity=26.99556]
44: 16745041 Pavel_Buran

```
[pem_pos=0, final_scores=0.54139197,
final_local_scores=0.6882469, pem=0.0, similarity=24.628374]
45: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.57249373,
final_local_scores=0.562086, pem=-0.27794555,
similarity=22.37499]
46: 19109147 Michael_Hübner
    [pem_pos=0, final_scores=0.6311027,
final_local_scores=0.7832014, pem=0.0, similarity=27.462006]
47: 11867 Germany
    [pem_pos=0, final_scores=0.8730441,
final_local_scores=0.87160283, pem=-0.15529114,
similarity=30.944576]
48: 16732280 Laurent_Gané
    [pem_pos=0, final_scores=0.48305416,
final_local_scores=0.64355576, pem=0.0, similarity=23.2947]
49: 5843419 France
    [pem_pos=0, final_scores=0.74041957,
final_local_scores=0.73239404, pem=-0.16240834,
similarity=26.829014]
50: 11438203 Marco_Villa
    [pem_pos=0, final_scores=0.36772305,
final_local_scores=0.489187, pem=0.0, similarity=18.688028]
51: 14532 Italy
    [pem_pos=0, final_scores=0.8130725,
final_local_scores=0.82617295, pem=-0.20612453,
similarity=29.8653]
52: 22665655 Scott_McGrory
    [pem_pos=0, final_scores=0.43855488,
final_local_scores=0.54411983, pem=0.0, similarity=20.327333]
53: 23728468 Stephen_Pate
    [pem_pos=0, final_scores=0.3794046,
final_local_scores=0.518312, pem=0.0, similarity=19.557177]
54: 4689264 Australia
    [pem_pos=0, final_scores=0.7456904,
final_local_scores=0.75835496, pem=-0.17168477,
similarity=27.654186]
55: 11003628 Andreas_Kappes
    [pem_pos=0, final_scores=0.5444307,
final_local_scores=0.6721311, pem=0.0, similarity=24.147446]
56: 14671526 Carsten_Wolf
    [pem_pos=0, final_scores=0.4328962,
final_local_scores=0.5877191, pem=0.0, similarity=21.628422]
57: 11867 Germany
    [pem_pos=0, final_scores=0.83885294,
final_local_scores=0.8342544, pem=-0.15529114,
similarity=29.830025]
58: 20417509 Kurt_Betschart
    [pem_pos=0, final_scores=0.5462552,
final_local_scores=0.68172836, pem=0.0, similarity=24.433847]
59: 15169629 Bruno_Risi
    [pem_pos=0, final_scores=0.5843898,
```

```
final_local_scores=0.73290616, pem=0.0, similarity=25.961094]
60: 26748 Switzerland
    [pem_pos=0, final_scores=0.59644324,
final_local_scores=0.5764383, pem=-0.18598008,
similarity=22.30317]
61: 11435812 Gabriel_Curuchet
    [pem_pos=0, final_scores=0.58179605,
final_local_scores=0.71014005, pem=0.0, similarity=25.281708]
62: 11435903 Juan_Curuchet
    [pem_pos=0, final_scores=0.5053089,
final_local_scores=0.652645, pem=0.0, similarity=23.56594]
63: 18951905 Argentina
    [pem_pos=0, final_scores=0.5795544,
final_local_scores=0.57982594, pem=-0.18885723,
similarity=22.41991]
64: 21148 Netherlands
    [pem_pos=0, final_scores=0.6491626,
final_local_scores=0.64168084, pem=-0.24940795,
similarity=24.595068]
65: 76972 Denmark
    [pem_pos=0, final_scores=0.616885,
final_local_scores=0.6164125, pem=-0.26762357,
similarity=23.940071]
66: 12935832 Joan_Llaneras
    [pem_pos=0, final_scores=0.42116398,
final_local_scores=0.55905247, pem=0.0, similarity=20.772953]
67: 26667 Spain
    [pem_pos=0, final_scores=0.63677114,
final_local_scores=0.62824315, pem=-0.18857051,
similarity=23.863216]
68: 26964606 Austria
    [pem_pos=0, final_scores=0.53757936,
final_local_scores=0.5071796, pem=-0.25262418,
similarity=20.598774]
69: 5843419 France
    [pem_pos=0, final_scores=0.47577488,
final_local_scores=0.4433099, pem=-0.16240834,
similarity=18.202166]
```

Sample: b'643_CYCLING'

CYCLING - [1 1 BOARDMAN 1] FULFILS WORLD RECORD PREDICTION .
Martin Ayres [2 2 MANCHESTER 2] , [3 3 England 3] 1996-08-28 [
4 4 Britain 4] 's [5 5 Chris Boardman 5] fulfilled his
prediction that he would break the world 4,000 metres world
record in the world track cycling championships on
Wednesday . [6 6 Boardman 6] clocked four minutes 13.353
seconds to slice more than six seconds from the previous
world mark of 4:19.699 set by 1 Olympic champion [7 7 Andrea C
ollinelli 7] of [8 8 Italy 8] in [9 9 Atlanta 9] in July . [10
10 Collinelli 10] qualified in second place , also beating
his old record in 4:17.696 . " I was very nervous before the
start but then I was amazed by the speed of my ride , " [11 11

Boardman 11] said . " I do n't know if I can go any faster . Who knows what will happen in the later stages ? " [12 12 Boardman 12] , 28 , did not contest the 2 Olympic pursuit because of its proximity to the [13 13 Tour de France 13] in which he led the [14 14 French 14] 3 GAN team . However , he took a bronze medal in the 4 Olympic road time-trial and then returned home to prepare for the world track series in the [15 15 Manchester 15] indoor velodrome . He adopted the " superman " riding position with arms at full stretch perfected by fellow [16 16 Briton 16] [17 17 Graeme Obree 17] , the 1995 world champion , and taken up in [18 18 Atlanta 18] by [19 19 Collinelli 19] . [20 20 Obree 20] was forced to pull out of his title defence because of a viral infection . Qualifiers for Wednesday evening 's quarter-finals : 1. [21 21 Chris Boardman 21] ([22 22 Britain 22]) 4:13.353 2. [23 23 Andrea Collinelli 23] ([24 24 Italy 24]) 4:17.696 3. Mariano Friedick ([25 25 U.S. 25]) 4:19.808 4. Heiko Szonn ([26 26 Germany 26]) 4:21.009 5. [27 27 Francis Moreau 27] ([28 28 France 28]) 4:21.454 6. [29 29 Alexei Markov 29] ([30 30 Russia 30]) 4:22.738 7. Michael Sandstod ([31 31 Denmark 31]) 4:24.427 8. [32 32 Edouard Gritsoun 32] ([33 33 Russia 33]) 4:26.467

1: 917770 Chris_Boardman
2: 20206 Manchester
3: 9316 England
4: 31717 United_Kingdom
5: 917770 Chris_Boardman
6: 917770 Chris_Boardman
7: 17826490 Andrea_Collinelli
8: 14532 Italy
9: 3138 Atlanta
10: 17826490 Andrea_Collinelli
11: 917770 Chris_Boardman
12: 917770 Chris_Boardman
13: 30498 Tour_de_France
14: 5843419 France
15: 20206 Manchester
16: 31717 United_Kingdom
17: 1210552 Graeme_Obree
18: 3138 Atlanta
19: 17826490 Andrea_Collinelli
20: 1210552 Graeme_Obree
21: 917770 Chris_Boardman
22: 31717 United_Kingdom
23: 17826490 Andrea_Collinelli
24: 14532 Italy
25: 3434750 United_States
26: 11867 Germany
27: 18894901 Francis_Moreau
28: 5843419 France
29: 12752327 Alexei_Markov
30: 25391 Russia

31: 76972 Denmark
32: 24395210 Eduard_Gritsun
33: 25391 Russia
1: 72311 1996_Summer_Olympics
[winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.24183507, final_local_scores=-0.29457802,
pem=-3.2532303, similarity=12.99049]
RECALL MISS: 72311 1996_Summer_Olympics]
2: 72311 1996_Summer_Olympics
[winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.1923337, final_local_scores=-0.24050528,
pem=-3.2532303, similarity=14.604129]
RECALL MISS: 72311 1996_Summer_Olympics]
3: 1258234 Crédit_Agricole_(cycling_team)
[winner: 2363282 Gan_(Laamu_Atoll), pem_pos=0,
final_scores=-0.5584639, final_local_scores=-0.43678388,
pem=-0.73719966, similarity=-4.935783]
RECALL MISS: 1258234 Crédit_Agricole_(cycling_team)]
4: 72311 1996_Summer_Olympics
[winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.27808607, final_local_scores=-0.33417672,
pem=-3.2532303, similarity=11.808786]
RECALL MISS: 72311 1996_Summer_Olympics]
1: 917770 Chris_Boardman
[pem_pos=0, final_scores=1.4247962,
final_local_scores=1.5179675, pem=0.0, similarity=49.388893]
2: 20206 Manchester
[pem_pos=0, final_scores=0.5122475,
final_local_scores=0.5194618, pem=-0.3209542,
similarity=21.33689]
3: 9316 England
[pem_pos=0, final_scores=0.6327418,
final_local_scores=0.6644411, pem=-0.19053856,
similarity=24.95414]
4: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5996093,
final_local_scores=0.61420757, pem=-0.61114657,
similarity=25.742403]
5: 917770 Chris_Boardman
[pem_pos=0, final_scores=0.7431957,
final_local_scores=0.767404, pem=0.0, similarity=26.99058]
6: 917770 Chris_Boardman
[pem_pos=0, final_scores=1.5630734,
final_local_scores=1.6690142, pem=0.0, similarity=53.896427]
7: 17826490 Andrea_Collinelli
[pem_pos=0, final_scores=0.55123717,
final_local_scores=0.65546024, pem=0.0, similarity=23.649954]
8: 14532 Italy
[pem_pos=0, final_scores=0.70672673,
final_local_scores=0.74914414, pem=-0.20612453,
similarity=27.566607]
9: 3138 Atlanta

```
[pem_pos=0, final_scores=0.44478232,
final_local_scores=0.45586652, pem=-0.23152597,
similarity=18.952753]
10: 17826490 Andrea_Collinelli
    [pem_pos=0, final_scores=0.70290726,
final_local_scores=0.8211368, pem=0.0, similarity=28.594072]
11: 917770 Chris_Boardman
    [pem_pos=0, final_scores=1.0888835,
final_local_scores=1.1510339, pem=0.0, similarity=38.438858]
12: 917770 Chris_Boardman
    [pem_pos=0, final_scores=1.5372963,
final_local_scores=1.6408567, pem=0.0, similarity=53.056152]
13: 30498 Tour_de_France
    [pem_pos=0, final_scores=0.5508746,
final_local_scores=0.5733092, pem=-0.1843033,
similarity=22.200672]
14: 5843419 France
    [pem_pos=0, final_scores=0.7272677,
final_local_scores=0.75577265, pem=-0.80223274,
similarity=31.006142]
15: 20206 Manchester
    [pem_pos=0, final_scores=0.49319047,
final_local_scores=0.49864483, pem=-0.3209542,
similarity=20.715668]
16: 31717 United_Kingdom
    [pem_pos=1, final_scores=0.37404072,
final_local_scores=0.36780816, pem=-0.9694005,
similarity=20.337593]
17: 1210552 Graeme_Obree
    [pem_pos=0, final_scores=0.63670343,
final_local_scores=0.70707834, pem=0.0, similarity=25.19034]
18: 3138 Atlanta
    [pem_pos=0, final_scores=0.5615888,
final_local_scores=0.5834599, pem=-0.23152597,
similarity=22.760393]
19: 17826490 Andrea_Collinelli
    [pem_pos=0, final_scores=0.53077,
final_local_scores=0.6331031, pem=0.0, similarity=22.98277]
20: 1210552 Graeme_Obree
    [pem_pos=0, final_scores=0.5205266,
final_local_scores=0.5801728, pem=0.0, similarity=21.403225]
21: 917770 Chris_Boardman
    [pem_pos=0, final_scores=0.8461738,
final_local_scores=0.87989193, pem=0.0, similarity=30.347445]
22: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.58825016,
final_local_scores=0.6017994, pem=-0.61114657,
similarity=25.37212]
23: 17826490 Andrea_Collinelli
    [pem_pos=0, final_scores=0.60209507,
final_local_scores=0.7110148, pem=0.0, similarity=25.307814]
24: 14532 Italy
```

```
[pem_pos=0, final_scores=0.7102416,
final_local_scores=0.7529836, pem=-0.20612453,
similarity=27.681185]
25: 3434750 United_States
    [pem_pos=0, final_scores=0.4948477,
final_local_scores=0.511298, pem=-0.14075686,
similarity=20.113323]
26: 11867 Germany
    [pem_pos=0, final_scores=0.75120586,
final_local_scores=0.79078776, pem=-0.15529114,
similarity=28.532894]
27: 18894901 Francis_Moreau
    [pem_pos=0, final_scores=0.637476,
final_local_scores=0.7455387, pem=0.0, similarity=26.338076]
28: 5843419 France
    [pem_pos=0, final_scores=0.7088904,
final_local_scores=0.73569834, pem=-0.16240834,
similarity=26.92762]
29: 12752327 Alexei_Markov
    [pem_pos=0, final_scores=0.51485974,
final_local_scores=0.6665906, pem=0.0, similarity=23.982105]
30: 25391 Russia
    [pem_pos=0, final_scores=0.77025586,
final_local_scores=0.852434, pem=-0.20299491,
similarity=30.63196]
31: 76972 Denmark
    [pem_pos=0, final_scores=0.5613637,
final_local_scores=0.60465074, pem=-0.26762357,
similarity=23.589075]
32: 24395210 Eduard_Gritsun
    [pem_pos=0, final_scores=0.57864475,
final_local_scores=0.7557743, pem=0.0, similarity=26.643526]
33: 25391 Russia
    [pem_pos=0, final_scores=0.7360764,
final_local_scores=0.81509805, pem=-0.20299491,
similarity=29.517782]
```

Sample: b'644_CYCLING'
CYCLING - [1 1 BOARDMAN 1] BREAKS WORLD 4,000 METRES
RECORD . [2 2 MANCHESTER 2] , [3 3 England 3] 1996-08-28 [4 4
Britain 4] 's [5 5 Chris Boardman 5] broke the world 4,000
metres cycling record by more than six seconds at the world
championships on Wednesday . [6 6 Boardman 6] clocked four
minutes 13.353 seconds in the qualifying round of the
individual pursuit event to beat the previous mark of
4:19.699 set by [7 7 Andrea Collinelli 7] of [8 8 Italy 8] at
the [9 9 Atlanta 9] 1 Olympics on July 24 .

```
1: 917770 Chris_Boardman
2: 20206 Manchester
3: 9316 England
4: 31717 United_Kingdom
5: 917770 Chris_Boardman
```

```

6: 917770 Chris_Boardman
7: 17826490 Andrea_Collinelli
8: 14532 Italy
9: 3138 Atlanta
1: 72311 1996_Summer_Olympics
    [winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.043946087, final_local_scores=-0.064762235,
pem=-2.7333555, similarity=17.021492]
    gt: 72311 1996_Summer_Olympics, pem_pos=14,
final_scores=-0.1313884, final_local_scores=-0.14608741,
pem=-4.001833, similarity=21.492764 ]
1: 917770 Chris_Boardman
    [pem_pos=0, final_scores=1.3527317,
final_local_scores=1.430872, pem=0.0, similarity=46.789787]
2: 20206 Manchester
    [pem_pos=0, final_scores=0.5438917,
final_local_scores=0.56442255, pem=-0.3209542,
similarity=22.678606]
3: 9316 England
    [pem_pos=0, final_scores=0.61907506,
final_local_scores=0.6721417, pem=-0.19053856,
similarity=25.183939]
4: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5684084,
final_local_scores=0.6119314, pem=-0.61114657,
similarity=25.674475]
5: 917770 Chris_Boardman
    [pem_pos=0, final_scores=0.931011,
final_local_scores=0.9524078, pem=0.0, similarity=32.51146]
6: 917770 Chris_Boardman
    [pem_pos=0, final_scores=1.6810818,
final_local_scores=1.7895446, pem=0.0, similarity=57.493294]
7: 17826490 Andrea_Collinelli
    [pem_pos=0, final_scores=0.54288894,
final_local_scores=0.6906012, pem=0.0, similarity=24.698631]
8: 14532 Italy
    [pem_pos=0, final_scores=0.6376647,
final_local_scores=0.72026706, pem=-0.20612453,
similarity=26.704857]
9: 3138 Atlanta
    [pem_pos=0, final_scores=0.39209038,
final_local_scores=0.4224723, pem=-0.23152597,
similarity=17.956202]

```

Sample: b'645_SOCCER'

```

SOCCER - [1 1 ENGLISH 1] FIRST DIVISION RESULTS /
STANDINGS . [2 2 LONDON 2] 1996-08-28 Results of [3 3 English
3] first division soccer matches on Wednesday : [4 4 Barnsley
4] 3 [5 5 Reading 5] 0 [6 6 Stoke 6] 1 1 Bradford 0 [7 7 Swin
don 7] 1 [8 8 Oldham 8] 0 [9 9 Wolverhampton 9] 1 [10 10 Quee
ns Park Rangers 10] 1 Standings ( tabulated under played ,
won , drawn , lost , goals for , against , points ) : [11 11 B

```

arnsley 11] 3 3 0 0 8 2 9 [12 12 Stoke 12] 3 3 0 0 5 2 9 [13 1
3 Tranmere 13] 3 2 1 0 6 3 7 [14 14 Bolton 14] 3 2 1 0 5 2 7 [
15 15 Wolverhampton 15] 3 2 1 0 5 2 7 [16 16 Queens Park Rang
ers 16] 3 2 1 0 5 3 7 [17 17 Norwich 17] 3 2 0 1 4 3 6 [18 18
Ipswich 18] 3 1 1 1 6 4 4 [19 19 Birmingham 19] 2 1 1 0 5 4 4
[20 20 Crystal Palace 20] 3 1 1 1 3 2 4 [21 21 Swindon 21] 3
1 1 1 2 3 4 [22 22 Oxford 22] 3 1 0 2 6 3 3 2 Bradford 3 1 0
2 3 3 3 [23 23 Huddersfield 23] 2 1 0 1 3 3 3 [24 24 Portsmou
th 24] 3 1 0 2 3 5 3 [25 25 Reading 25] 3 1 0 2 3 8 3 [26 26 M
an City 26] 3 1 0 2 2 3 3 [27 27 West Bromwich 27] 3 0 2 1 2
3 2 [28 28 Port Vale 28] 3 0 2 1 2 4 2 [29 29 Sheffield United
29] 2 0 1 1 4 5 1 [30 30 Grimsby 30] 3 0 1 2 4 7 1 [31 31 Cha
rlton 31] 2 0 1 1 1 3 1 [32 32 Southend 32] 3 0 1 2 1 7 1 [33
33 Oldham 33] 3 0 0 3 2 6 0

1: 9316 England
2: 17867 London
3: 9316 England
4: 333188 Barnsley_F.C.
5: 168007 Reading_F.C.
6: 203434 Stoke_City_F.C.
7: 74291 Swindon_Town_F.C.
8: 451131 Oldham_Athletic_A.F.C.
9: 217730 Wolverhampton_Wanderers_F.C.
10: 84588 Queens_Park_Rangers_F.C.
11: 333188 Barnsley_F.C.
12: 203434 Stoke_City_F.C.
13: 69219 Tranmere_Rovers_F.C.
14: 298601 Bolton_Wanderers_F.C.
15: 217730 Wolverhampton_Wanderers_F.C.
16: 84588 Queens_Park_Rangers_F.C.
17: 248083 Norwich_City_F.C.
18: 352843 Ipswich_Town_F.C.
19: 185196 Birmingham_City_F.C.
20: 385313 Crystal_Palace_F.C.
21: 74291 Swindon_Town_F.C.
22: 451165 Oxford_United_F.C.
23: 236211 Huddersfield_Town_F.C.
24: 564301 Portsmouth_F.C.
25: 168007 Reading_F.C.
26: 165813 Manchester_City_F.C.
27: 33921 West_Bromwich_Albiaon_F.C.
28: 203456 Port_Vale_F.C.
29: 314933 Sheffield_United_F.C.
30: 451129 Grimsby_Town_F.C.
31: 6720 Charlton_Athletic_F.C.
32: 451168 Southend_United_F.C.
33: 451131 Oldham_Athletic_A.F.C.
1: 431144 Bradford_City_A.F.C.
[winner: 1095011 Bradford_Bulls, pem_pos=9,
final_scores=-0.3860078, final_local_scores=-0.45383593,
pem=-3.818427, similarity=11.311543]
RECALL MISS: 431144 Bradford_City_A.F.C.]

2: 431144 Bradford_City_A.F.C.
[winner: 1095011 Bradford_Bulls, pem_pos=9,
final_scores=-0.32981646, final_local_scores=-0.39245546,
pem=-3.818427, similarity=13.143259]
RECALL MISS: 431144 Bradford_City_A.F.C.]
1: 9316 England
[pem_pos=1, final_scores=0.4996091,
final_local_scores=0.5569833, pem=-1.3527626,
similarity=28.06774]
2: 17867 London
[pem_pos=0, final_scores=0.6128172,
final_local_scores=0.69458354, pem=-0.09086737,
similarity=25.31162]
3: 9316 England
[pem_pos=1, final_scores=0.62524587,
final_local_scores=0.6942224, pem=-1.3527626,
similarity=32.163227]
4: 333188 Barnsley_F.C.
[pem_pos=0, final_scores=0.39359254,
final_local_scores=0.31503853, pem=-0.7126868,
similarity=17.366793]
5: 168007 Reading_F.C.
[pem_pos=0, final_scores=1.066872,
final_local_scores=1.0504906, pem=-1.2007197,
similarity=41.96815]
6: 203434 Stoke_City_F.C.
[pem_pos=0, final_scores=0.65809137,
final_local_scores=0.6071233, pem=-0.3417574,
similarity=24.066013]
7: 74291 Swindon_Town_F.C.
[pem_pos=4, final_scores=0.71485347,
final_local_scores=0.6700899, pem=-2.489671,
similarity=37.62575]
8: 451131 Oldham_Athletic_A.F.C.
[pem_pos=3, final_scores=0.8343796,
final_local_scores=0.7983667, pem=-2.8632221,
similarity=43.485214]
9: 217730 Wolverhampton_Wanderers_F.C.
[pem_pos=1, final_scores=0.89610606,
final_local_scores=0.8605921, pem=-2.274954,
similarity=42.14305]
10: 84588 Queens_Park_Rangers_F.C.
[pem_pos=0, final_scores=0.5038832,
final_local_scores=0.4377241, pem=0.0, similarity=17.152273]
11: 333188 Barnsley_F.C.
[pem_pos=0, final_scores=0.38877738,
final_local_scores=0.30977866, pem=-0.7126868,
similarity=17.209827]
12: 203434 Stoke_City_F.C.
[pem_pos=0, final_scores=0.5429604,
final_local_scores=0.4813602, pem=-0.3417574,
similarity=20.31299]

13: 69219 Tranmere_Rovers_F.C.
[pem_pos=0, final_scores=0.43716508,
final_local_scores=0.37243307, pem=-0.6002322,
similarity=18.468016]
14: 298601 Bolton_Wanderers_F.C.
[pem_pos=2, final_scores=0.565702,
final_local_scores=0.5020091, pem=-2.2845325,
similarity=31.49431]
15: 217730 Wolverhampton_Wanderers_F.C.
[pem_pos=1, final_scores=0.81314665,
final_local_scores=0.7699716, pem=-2.274954,
similarity=39.438755]
16: 84588 Queens_Park_Rangers_F.C.
[pem_pos=0, final_scores=0.5044718,
final_local_scores=0.438367, pem=0.0, similarity=17.17146]
17: 248083 Norwich_City_F.C.
[pem_pos=2, final_scores=0.85178906,
final_local_scores=0.81464213, pem=-2.7278075,
similarity=43.2345]
18: 352843 Ipswich_Town_F.C.
[pem_pos=4, final_scores=0.82768184,
final_local_scores=0.79425776, pem=-2.6931517,
similarity=42.437725]
19: 185196 Birmingham_City_F.C.
[pem_pos=2, final_scores=0.7745051,
final_local_scores=0.7311861, pem=-3.102209,
similarity=42.78006]
20: 385313 Crystal_Palace_F.C.
[pem_pos=0, final_scores=0.52662194,
final_local_scores=0.46079463, pem=-0.51090026,
similarity=20.619099]
21: 74291 Swindon_Town_F.C.
[pem_pos=4, final_scores=0.70507145,
final_local_scores=0.6594046, pem=-2.489671,
similarity=37.30688]
22: 451165 Oxford_United_F.C.
[pem_pos=18, final_scores=0.9884525,
final_local_scores=0.9734488, pem=-4.3872933,
similarity=56.99815]
23: 236211 Huddersfield_Town_F.C.
[pem_pos=3, final_scores=0.7270842,
final_local_scores=0.6800985, pem=-2.2859883,
similarity=36.81677]
24: 564301 Portsmouth_F.C.
[pem_pos=1, final_scores=0.7461042,
final_local_scores=0.70081407, pem=-1.2302613,
similarity=31.693752]
25: 168007 Reading_F.C.
[pem_pos=0, final_scores=0.9973164,
final_local_scores=0.9745117, pem=-1.2007197,
similarity=39.700787]
26: 165813 Manchester_City_F.C.

```

[pem_pos=0, final_scores=0.50765663,
final_local_scores=0.44975883, pem=0.0, similarity=17.511414]
27: 33921 West_Bromwich_Albion_F.C.
[pem_pos=2, final_scores=0.6090191,
final_local_scores=0.55191785, pem=-1.5002033,
similarity=28.71838]
28: 203456 Port_Vale_F.C.
[pem_pos=0, final_scores=0.587976,
final_local_scores=0.5412477, pem=0.0, similarity=20.241625]
29: 314933 Sheffield_United_F.C.
[pem_pos=0, final_scores=0.34907418,
final_local_scores=0.26565418, pem=0.0, similarity=12.017364]
30: 451129 Grimsby_Town_F.C.
[pem_pos=2, final_scores=0.8156927,
final_local_scores=0.77995104, pem=-2.6069987,
similarity=41.54227]
31: 6720 Charlton_Athletic_F.C.
[pem_pos=1, final_scores=0.78616315,
final_local_scores=0.7485215, pem=-2.103781,
similarity=37.867775]
32: 451168 Southend_United_F.C.
[pem_pos=2, final_scores=0.7275401,
final_local_scores=0.6834466, pem=-1.771178,
similarity=34.11706]
33: 451131 Oldham_Athletic_A.F.C.
[pem_pos=3, final_scores=0.695465,
final_local_scores=0.6466237, pem=-2.8632221,
similarity=38.9569]

```

Sample: b'646_CRICKET'

CRICKET - [1 1 ENGLISH 2 COUNTY CHAMPIONSHIP 1] SCORES . [1 1 LONDON 1] 1996-08-28 Close of play scores on the first day of four-day [2 2 English 2] [3 3 County Championship 3] cricket matches on Wednesday : At [4 4 Portsmouth 4] : [5 5 Middlesex 5] 199 in 60 overs (K. Brown 57 ; L. Botham 5-67) . [6 6 Hampshire 6] 105-4 . At [7 7 Chester-le-Street 7] : [8 8 Glamorgan 8] 73-3 v [9 9 Durham 9] .

1: 17867 London
2: 9316 England
3: 683599 County_Championship
4: 24207 Portsmouth
5: 822103 Middlesex_County_Cricket_Club
6: 1919370 Hampshire_County_Cricket_Club
7: 293089 Chester-le-Street
8: 1622320 Glamorgan_County_Cricket_Club
9: 1176371 Durham_County_Cricket_Club
1: 9316 England
[winner: 683599 County_Championship, pem_pos=0,
final_scores=0.2932747, final_local_scores=0.2320647,
pem=0.0, similarity=11.014986]
RECALL MISS: 9316 England]

2: 683599 County_Championship

```
[winner: 683599 County_Championship, pem_pos=0,
final_scores=0.2932747, final_local_scores=0.2320647,
pem=0.0, similarity=11.014986]
  gt: 683599 County_Championship, pem_pos=0,
final_scores=0.2932747, final_local_scores=0.2320647,
pem=0.0, similarity=11.014986 ]
1: 17867 London
  [pem_pos=0, final_scores=0.6796939,
final_local_scores=0.7417883, pem=-0.09086737,
similarity=26.720306]
2: 9316 England
  [pem_pos=1, final_scores=0.2793222,
final_local_scores=0.2731566, pem=-1.3527626,
similarity=19.597784]
3: 683599 County_Championship
  [pem_pos=0, final_scores=0.19581847,
final_local_scores=0.1256086, pem=0.0, similarity=7.838124]
4: 24207 Portsmouth
  [pem_pos=0, final_scores=0.38869327,
final_local_scores=0.41641843, pem=-0.75559247,
similarity=20.6255]
5: 822103 Middlesex_County_Cricket_Club
  [pem_pos=1, final_scores=0.82540876,
final_local_scores=0.80562556, pem=-1.5695794,
similarity=36.6668]
6: 1919370 Hampshire_County_Cricket_Club
  [pem_pos=1, final_scores=0.6874614,
final_local_scores=0.6637637, pem=-1.9232318,
similarity=34.35658]
7: 293089 Chester-le-Street
  [pem_pos=0, final_scores=0.36669284,
final_local_scores=0.3886788, pem=-0.03024954,
similarity=15.853165]
8: 1622320 Glamorgan_County_Cricket_Club
  [pem_pos=1, final_scores=0.5174734,
final_local_scores=0.47551572, pem=-0.7903395,
similarity=22.57804]
9: 1176371 Durham_County_Cricket_Club
  [pem_pos=2, final_scores=0.6123386,
final_local_scores=0.5920866, pem=-2.3952909,
similarity=34.784725]
1: 683599 County_Championship
  [pem_pos=0, final_scores=0.2932747,
final_local_scores=0.2320647, pem=0.0, similarity=11.014986]
```

Sample: b'647_SOCCER'

SOCCER - [1 1 FOWLER 1] AND [2 2 McMANAMAN 2] OUT OF [3 3 EN
GLAND 3] SQUAD . [4 4 LONDON 4] 1996-08-28 [5 5 England 5]
soccer manager [6 6 Glen Hoddle 6] confirmed on Wednesday
that the [7 7 Liverpool 7] pair of [8 8 Steve McManaman 8]
and [9 9 Robbie Fowler 9] would miss [10 10 England 10] 's [1
1 World Cup 1] qualifying match against [11 11 Moldova 11] on

Sunday . The two men , both suffering from back injuries , joined the [12 12 England 12] squad at training but it was soon clear they had no chance of making the flight to [13 13 K ishinev 13] on Friday . " They had some scans and X-rays yesterday and they 're out , " said [14 14 Hoddle 14] . " The more important thing for me was to get them down here and have a chat . To go another five weeks without that chance would have been foolish , " he added . [15 15 Hoddle 15] , who has already lost the services of midfielder [16 16 Darren Anderton 16] and defender [17 17 Steve Howey 17] , delayed naming any replacements . There was also concern over injuries to [18 18 Paul Gascoigne 18] , [19 19 Les Ferdinand 19] and [20 20 David Batty 20] .

1: 346075 Robbie_Fowler
2: 1412864 Steve_McManaman
3: 9904 England_national_football_team
4: 17867 London
5: 9904 England_national_football_team
6: 552843 Glenn_Hoddle
7: 18119 Liverpool_F.C.
8: 1412864 Steve_McManaman
9: 346075 Robbie_Fowler
10: 9904 England_national_football_team
11: 1011292 Moldova_national_football_team
12: 9904 England_national_football_team
13: 56635 Chișinău
14: 552843 Glenn_Hoddle
15: 552843 Glenn_Hoddle
16: 1089164 Darren_Anderton
17: 2201395 Steve_Howey_(footballer)
18: 413242 Paul_Gascoigne
19: 526001 Les_Ferdinand
20: 1421688 David_Batty
1: 11370 FIFA_World_Cup
 [winner: 126460 1998_FIFA_World_Cup, pem_pos=18,
final_scores=0.23976134, final_local_scores=0.21734297,
pem=-4.2172847, similarity=33.50991]
 gt: 11370 FIFA_World_Cup, pem_pos=0,
final_scores=0.09387474, final_local_scores=0.06340201,
pem=-1.001087, similarity=11.425822]
1: 346075 Robbie_Fowler
 [pem_pos=0, final_scores=1.1587901,
final_local_scores=1.1882384, pem=0.0, similarity=39.549114]
2: 1412864 Steve_McManaman
 [pem_pos=0, final_scores=1.0512145,
final_local_scores=1.0703554, pem=0.0, similarity=36.031254]
3: 9904 England_national_football_team
 [pem_pos=1, final_scores=0.42783374,
final_local_scores=0.3658719, pem=-2.776464,
similarity=30.106903]
4: 17867 London
 [pem_pos=0, final_scores=0.49080616,

```
final_local_scores=0.5713469, pem=-0.09086737,
similarity=21.633993]
5: 9904 England_national_football_team
    [pem_pos=1, final_scores=0.4209618,
final_local_scores=0.3583653, pem=-2.776464,
similarity=29.882893]
6: 552843 Glenn_Hoddle
    [pem_pos=0, final_scores=0.76184326,
final_local_scores=0.7456448, pem=0.0, similarity=26.341242]
7: 18119 Liverpool_F.C.
    [pem_pos=1, final_scores=0.6394158,
final_local_scores=0.60258234, pem=-1.1702383,
similarity=28.43591]
8: 1412864 Steve_McManaman
    [pem_pos=0, final_scores=0.76935226,
final_local_scores=0.7624638, pem=0.0, similarity=26.843153]
9: 346075 Robbie_Fowler
    [pem_pos=0, final_scores=0.92469317,
final_local_scores=0.9275238, pem=0.0, similarity=31.768871]
10: 9904 England_national_football_team
    [pem_pos=1, final_scores=0.45170325,
final_local_scores=0.39194572, pem=-2.776464,
similarity=30.884996]
11: 1011292 Moldova_national_football_team
    [pem_pos=1, final_scores=0.57431287,
final_local_scores=0.6256529, pem=-2.4170835,
similarity=35.90492]
12: 9904 England_national_football_team
    [pem_pos=1, final_scores=0.424074,
final_local_scores=0.3617649, pem=-2.776464,
similarity=29.984343]
13: 56635 Chișinău
    [pem_pos=0, final_scores=0.39166725,
final_local_scores=0.52286243, pem=0.0, similarity=19.69297]
14: 552843 Glenn_Hoddle
    [pem_pos=0, final_scores=0.49729794,
final_local_scores=0.4566692, pem=0.0, similarity=17.717632]
15: 552843 Glenn_Hoddle
    [pem_pos=0, final_scores=0.66199183,
final_local_scores=0.63657224, pem=0.0, similarity=23.086298]
16: 1089164 Darren_Anderton
    [pem_pos=0, final_scores=0.7261962,
final_local_scores=0.75022346, pem=0.0, similarity=26.477877]
17: 2201395 Steve_Howey_(footballer)
    [pem_pos=1, final_scores=0.5805606,
final_local_scores=0.6236265, pem=-0.3800337,
similarity=24.766653]
18: 413242 Paul_Gascoigne
    [pem_pos=0, final_scores=0.7273579,
final_local_scores=0.7188923, pem=0.0, similarity=25.542892]
19: 526001 Les_Ferdinand
    [pem_pos=0, final_scores=0.6641254,
```

```
final_local_scores=0.66374916, pem=0.0, similarity=23.897312]
20: 1421688 David_Batty
    [pem_pos=0, final_scores=0.5230928,
final_local_scores=0.53801066, pem=0.0, similarity=20.145023]
1: 126460 1998_FIFA_World_Cup
    [pem_pos=18, final_scores=0.23976134,
final_local_scores=0.21734297, pem=-4.2172847,
similarity=33.50991]
```

Sample: b'648_HORSE'

HORSE RACING - JOCKEY WEAVER RECEIVES 21-DAY BAN . [1 1 LOND
ON 1] 1996-08-28 Leading rider Jason Weaver received a 21-day
ban from the disciplinary committee of the Jockey Club on
Wednesday . Weaver had been reported after being found guilty
of irresponsible riding at the provincial track of [2 2 Ponte
fract 2] 10 days ago . It was his fourth riding offence this
season . Although five days of the ban were suspended until
January 1 , Weaver will miss next month 's big [3 3 St Leger 3
] meeting , including the ride on top stayer Double Trigger
in the [4 4 Doncaster Cup 4] . Weaver shot to prominence in
1994 when he won the [5 5 English 5] [6 6 2,000 Guineas 6] on
Mister Baileys in his first ride in a classic .

```
1: 17867 London
2: 152120 Pontefract
3: 68032 St._Leger_Stakes
4: 742031 Doncaster_Cup
5: 9316 England
6: 68030 2,000_Guineas_Stakes
1: 17867 London
    [pem_pos=0, final_scores=0.5827525,
final_local_scores=0.63961226, pem=-0.09086737,
similarity=23.67117]
2: 152120 Pontefract
    [pem_pos=0, final_scores=0.24964903,
final_local_scores=0.26988816, pem=-0.05904496,
similarity=12.464809]
3: 68032 St._Leger_Stakes
    [pem_pos=0, final_scores=0.4738428,
final_local_scores=0.47887498, pem=0.0, similarity=18.380297]
4: 742031 Doncaster_Cup
    [pem_pos=0, final_scores=0.47506368,
final_local_scores=0.51143646, pem=0.0, similarity=19.351997]
5: 9316 England
    [pem_pos=1, final_scores=0.5091504,
final_local_scores=0.5327932, pem=-1.3527626,
similarity=27.34586]
6: 68030 2,000_Guineas_Stakes
    [pem_pos=0, final_scores=0.24318843,
final_local_scores=0.2415266, pem=0.0, similarity=11.297349]
```

Sample: b'649_ROWING'

ROWING - [1 1 REDGRAVE 1] MAY SEEK FIFTH [1 1 OLYMPIC 1]

GOLD . [2 2 LONDON 2] 1996-08-28 [3 3 Britain 3] 's [4 4 Steven Redgrave 4] said on Wednesday he might change his mind and go for a fifth consecutive [5 5 Olympic 5] gold medal at the [6 6 2000 Games 6] in [7 7 Sydney 7] . [8 8 Redgrave 8] is one of only five athletes who have won gold medals at four successive [9 9 Olympics 9] . He shared victory with [10 10 Matthew Pinsent 10] in the coxless pairs at the [11 11 Atlanta Games 11] and said at the time that would be his last shot . But he has had second thoughts since then . " I 'm only 34 . Some say that 's too old for an athlete , " he said . " But I 'll be 38 by [12 12 Sydney 12] and that 's not too old . It 's whether I have got the enthusiasm for the training over the next four years . Rowing is an endurance sport . "

1: 201245 Steve_Redgrave
2: 17867 London
3: 31717 United_Kingdom
4: 201245 Steve_Redgrave
5: 27566 Summer_Olympic_Games
6: 71432 2000_Summer_Olympics
7: 27862 Sydney
8: 201245 Steve_Redgrave
9: 27566 Summer_Olympic_Games
10: 219303 Matthew_Pinsent
11: 72311 1996_Summer_Olympics
12: 71432 2000_Summer_Olympics
1: 71432 2000_Summer_Olympics
[winner: 22576 Olympic_Games, pem_pos=0,
final_scores=0.07218425, final_local_scores=-0.005453855,
pem=-0.38020664, similarity=5.9945817]
gt: 71432 2000_Summer_Olympics, pem_pos=17,
final_scores=-0.28754002, final_local_scores=-0.3861063,
pem=-4.537967, similarity=17.2457]
1: 201245 Steve_Redgrave
[pem_pos=0, final_scores=0.9297542,
final_local_scores=0.9566838, pem=0.0, similarity=32.639065]
2: 17867 London
[pem_pos=0, final_scores=0.651336,
final_local_scores=0.6806842, pem=-0.09086737,
similarity=24.896837]
3: 31717 United_Kingdom
[pem_pos=0, final_scores=0.48837996,
final_local_scores=0.5404503, pem=-0.61114657,
similarity=23.541338]
4: 201245 Steve_Redgrave
[pem_pos=0, final_scores=0.59897697,
final_local_scores=0.58421993, pem=0.0, similarity=21.524]
5: 27566 Summer_Olympic_Games
[pem_pos=2, final_scores=0.39375174,
final_local_scores=0.37089282, pem=-3.2532303,
similarity=32.84947]
6: 71432 2000_Summer_Olympics
[pem_pos=0, final_scores=0.26000816,

```
final_local_scores=0.2219229, pem=0.0, similarity=10.712335]
7: 27862 Sydney
    [pem_pos=0, final_scores=0.5463534,
final_local_scores=0.54888666, pem=-0.1443489,
similarity=21.254578]
8: 201245 Steve_Redgrave
    [pem_pos=0, final_scores=1.05812,
final_local_scores=1.096904, pem=0.0, similarity=36.823517]
9: 27566 Summer_Olympic_Games
    [pem_pos=2, final_scores=0.44305974,
final_local_scores=0.42475432, pem=-2.7333555,
similarity=31.62964]
10: 219303 Matthew_Pinsent
    [pem_pos=0, final_scores=0.6273089,
final_local_scores=0.6531683, pem=0.0, similarity=23.581558]
11: 72311 1996_Summer_Olympics
    [pem_pos=0, final_scores=0.3628661,
final_local_scores=0.34274116, pem=0.0, similarity=14.317793]
12: 71432 2000_Summer_Olympics
    [pem_pos=12, final_scores=0.24291568,
final_local_scores=0.20325196, pem=-5.244351,
similarity=38.67475]
1: 22576 Olympic_Games
    [pem_pos=0, final_scores=0.07218425,
final_local_scores=-0.005453855, pem=-0.38020664,
similarity=5.9945817]
```

Sample: b'650_CRICKET'

CRICKET - [1 1 BOTHAM 1] DISMISSES [2 2 GATTING 2] IN FIRST CLASS DEBUT . [3 3 LONDON 3] 1996-08-28 [4 4 Liam Botham 4] demonstrated his father 1 Ian 's golden touch on Wednesday shortly after making his county debut for [5 5 Hampshire 5] . [6 6 Botham 6] dismissed [7 7 Mike Gatting 7] with his seventh ball when the former [8 8 England 8] captain pushed a half-volley to square-leg on the first day of the four-day match against [9 9 Middlesex 9] at [10 10 Portsmouth 10] . Earlier [11 11 Botham 11] arrived in [12 12 Portsmouth 12] from [13 13 Southampton 13] only to be told his services would not be required . He then drove 40 kms back to play for the second XI to learn that [14 14 John Stephenson 14] had dropped out of the [15 15 Middlesex 15] match in the meantime with a shoulder injury . [16 16 Botham 16] dashed back to [17 17 Portsmouth 17] and took the field as the thrid over began . [18 18 Ian Botham 18] began his test career in 1977 by dismissing [19 19 Australian 19] captain [20 20 Greg Chappell 20] with a long hop and went on to become his country 's most successful all-rounder ever with 5,200 runs , 383 wickets plus 120 catches in 102 tests .

```
1: 1789550 Liam_Botham
2: 413174 Mike_Gatting
3: 17867 London
4: 1789550 Liam_Botham
```

```
5: 1919370 Hampshire_County_Cricket_Club
6: 1789550 Liam_Botham
7: 413174 Mike_Gatting
8: 759125 England_cricket_team
9: 822103 Middlesex_County_Cricket_Club
10: 24207 Portsmouth
11: 1789550 Liam_Botham
12: 24207 Portsmouth
13: 7920751 Southampton
14: 2410399 John_Stephenson_(cricketer,_born_1965)
15: 822103 Middlesex_County_Cricket_Club
16: 1789550 Liam_Botham
17: 24207 Portsmouth
18: 15524 Ian_Botham
19: 374833 Australia_national_cricket_team
20: 413137 Greg_Chappell
1: 15524 Ian_Botham
    [winner: 151608 Ian_Chappell, pem_pos=9,
final_scores=-0.42905962, final_local_scores=-0.47960776,
pem=-3.8114343, similarity=10.504434]
    RECALL MISS: 15524 Ian_Botham]
1: 1789550 Liam_Botham
    [pem_pos=0, final_scores=0.63561004,
final_local_scores=0.6339786, pem=0.0, similarity=23.0089]
2: 413174 Mike_Gatting
    [pem_pos=0, final_scores=0.5110515,
final_local_scores=0.5269387, pem=0.0, similarity=19.814613]
3: 17867 London
    [pem_pos=0, final_scores=0.6987743,
final_local_scores=0.7763343, pem=-0.09086737,
similarity=27.751226]
4: 1789550 Liam_Botham
    [pem_pos=0, final_scores=0.6596037,
final_local_scores=0.65640455, pem=0.0, similarity=23.678135]
5: 1919370 Hampshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6135536,
final_local_scores=0.6237723, pem=-1.9232318,
similarity=33.163155]
6: 1789550 Liam_Botham
    [pem_pos=0, final_scores=0.6987481,
final_local_scores=0.7054141, pem=0.0, similarity=25.140676]
7: 413174 Mike_Gatting
    [pem_pos=0, final_scores=0.5642336,
final_local_scores=0.5767412, pem=0.0, similarity=21.300821]
8: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.80907387,
final_local_scores=0.805361, pem=-3.3054674,
similarity=46.098938]
9: 822103 Middlesex_County_Cricket_Club
    [pem_pos=1, final_scores=0.6241114,
final_local_scores=0.61952853, pem=-1.5695794,
similarity=31.113298]
```

```
10: 24207 Portsmouth
    [pem_pos=0, final_scores=0.34568167,
final_local_scores=0.3745339, pem=-0.75559247,
similarity=19.375582]
11: 1789550 Liam_Botham
    [pem_pos=0, final_scores=0.7009403,
final_local_scores=0.70780873, pem=0.0, similarity=25.212137]
12: 24207 Portsmouth
    [pem_pos=0, final_scores=0.34548497,
final_local_scores=0.37431902, pem=-0.75559247,
similarity=19.36917]
13: 7920751 Southampton
    [pem_pos=0, final_scores=0.46686846,
final_local_scores=0.49379468, pem=-0.6752677,
similarity=22.497742]
14: 2410399 John_Stephenson_(cricketer,_born_1965)
    [pem_pos=2, final_scores=0.31508803,
final_local_scores=0.3691364, pem=-1.9116861,
similarity=25.501526]
15: 822103 Middlesex_County_Cricket_Club
    [pem_pos=1, final_scores=0.6153698,
final_local_scores=0.6099796, pem=-1.5695794,
similarity=30.828337]
16: 1789550 Liam_Botham
    [pem_pos=0, final_scores=0.6664905,
final_local_scores=0.6701776, pem=0.0, similarity=24.08915]
17: 24207 Portsmouth
    [pem_pos=0, final_scores=0.41483456,
final_local_scores=0.45007288, pem=-0.75559247,
similarity=21.629816]
18: 15524 Ian_Botham
    [pem_pos=0, final_scores=0.7826757,
final_local_scores=0.80502415, pem=0.0, similarity=28.11324]
19: 374833 Australia_national_cricket_team
    [pem_pos=2, final_scores=0.6667481,
final_local_scores=0.66836274, pem=-3.6253178,
similarity=43.75003]
20: 413137 Greg_Chappell
    [pem_pos=0, final_scores=0.7463811,
final_local_scores=0.79226923, pem=0.0, similarity=27.732607]
```

Sample: b'651_RUGBY'

[1 1 RUGBY UNION 1] - [2 2 CARLING 2] LEFT OUT OF [3 3 ENGLAND 3] TRAINING SQUAD . [4 4 LONDON 4] 1996-08-28 Former [5 5 England 5] captain [6 6 Will Carling 6] along with [7 7 Jeremy Guscott 7] , [8 8 Rory Underwood 8] and [9 9 Dean Richards 9] have been left out of [10 10 England 10] 's first training squad of the season . The quartet , who possess 244 international caps between them , were also omitted from a summer training camp but will still be in contention when the northern international season starts later this year . " Their qualities are well known to the selectors and they

will , of course , be considered when the season gets underway , " the [11 11 Rugby Football Union 11] said in a statement on Wednesday .

1: 25405 Rugby_union
2: 531810 Will_Carling
3: 691024 England_national_rugby_union_team
4: 17867 London
5: 691024 England_national_rugby_union_team
6: 531810 Will_Carling
7: 775184 Jeremy_Guscott
8: 544435 Rory_Underwood
9: 17340321 Dean_Richards_(rugby_union)
10: 691024 England_national_rugby_union_team
11: 696501 Rugby_Football_Union
1: 25405 Rugby_union
[pem_pos=0, final_scores=0.47568613,
final_local_scores=0.48052406, pem=0.0, similarity=18.42951]
2: 531810 Will_Carling
[pem_pos=0, final_scores=0.973491,
final_local_scores=1.0092478, pem=0.0, similarity=34.20768]
3: 691024 England_national_rugby_union_team
[pem_pos=3, final_scores=0.3101014,
final_local_scores=0.24514347, pem=-3.7874734,
similarity=32.00215]
4: 17867 London
[pem_pos=0, final_scores=0.56393033,
final_local_scores=0.64605594, pem=-0.09086737,
similarity=23.86346]
5: 691024 England_national_rugby_union_team
[pem_pos=3, final_scores=0.32298666,
final_local_scores=0.25921863, pem=-3.7874734,
similarity=32.422184]
6: 531810 Will_Carling
[pem_pos=0, final_scores=0.7433422,
final_local_scores=0.7438423, pem=0.0, similarity=26.28745]
7: 775184 Jeremy_Guscott
[pem_pos=0, final_scores=0.6083562,
final_local_scores=0.6352171, pem=0.0, similarity=23.045856]
8: 544435 Rory_Underwood
[pem_pos=0, final_scores=0.46130133,
final_local_scores=0.4975322, pem=0.0, similarity=18.937065]
9: 17340321 Dean_Richards_(rugby_union)
[pem_pos=1, final_scores=0.4284112,
final_local_scores=0.45879441, pem=-0.35570362,
similarity=19.715425]
10: 691024 England_national_rugby_union_team
[pem_pos=3, final_scores=0.34811276,
final_local_scores=0.28666514, pem=-3.7874734,
similarity=33.24124]
11: 696501 Rugby_Football_Union
[pem_pos=0, final_scores=0.3260792,
final_local_scores=0.31038162, pem=0.0, similarity=13.352119]

Sample: b'652_BASEBALL'

BASEBALL - [1 1 DODGERS 1] WIN FIFTH STRAIGHT . [2 2 MONTREAL 2] 1996-08-28 [3 3 Hideo Nomo 3] allowed a run in seven innings for his fifth win in seven road starts and [4 4 Greg Gagne 4] capped a three-run fourth with a two-run homer as the [5 5 Los Angeles Dodgers 5] claimed a 5-1 victory the [6 6 Montreal Expos 6] on Tuesday . With their fifth straight win , the [7 7 Dodgers 7] moved a half-game ahead of the [8 8 Expos 8] at the top of the wild card hunt behind [9 9 Nomo 9] (13-10) , who allowed six hits and walked four with six strikeouts . In [10 10 San Francisco 10] , [11 11 Mike Williams 11] allowed two runs in 7-1/3 innings and [12 12 Benito Santiago 12] and Ruben Amaro had RBI hits in the first inning as the [13 13 Philadelphia Phillies 13] edged the [14 14 San Francisco Giants 14] 3-2 . [15 15 Williams 15] (5-12) , who snapped a personal three-game losing streak , allowed five hits , walked two and struck out five . It was also [16 16 Williams 16] ' first win in three career decisions against [17 17 San Francisco 17] . In [18 18 Pittsburgh 18] , [19 19 Al Martin 19] 's run-scoring single snapped a fifth-inning tie and [20 20 Denny Neagle 20] outdueled [21 21 John Smoltz 21] as the [22 22 Pittsburgh Pirates 22] edged the [23 23 Atlanta Braves 23] 3-2 . The [24 24 Braves 24] led 2-1 entering the fifth , but the [25 25 Pirates 25] pushed across two runs against [26 26 Smoltz 26] (20-7) . [27 27 Neagle 27] (14-6) beat the [28 28 Braves 28] for the third time this season , allowing two runs and six hits in eight innings . In [29 29 St Louis 29] , [30 30 Gary Sheffield 30] and [31 31 Devon White 31] each drove in two runs and [32 32 Mark Hutton 32] scattered four hits over six innings to lead the [33 33 Florida Marlins 33] past the [34 34 St. Louis Cardinals 34] 6-3 . White added a solo homer , his 11th , off reliever [35 35 Mark Petkovsek 35] with one out in the fifth , giving the [36 36 Marlins 36] a 6-0 lead . In [37 37 New York 37] , [38 38 Steve Finley 38] 's three-run homer capped a four-run eighth inning and gave the [39 39 San Diego Padres 39] a 4-3 victory over [1 1 New York 1] , spoiling [40 40 Bobby Valentine 40] 's debut as [41 41 Mets 41] ' manager . The rally made a winner out of reliever Willie Blair Tony Gwynn and [42 42 Wally Joyner 42] had two hits apiece , helping the [43 43 Padres 43] to their third straight win . First-place [44 44 San Diego 44] has won seven of its last eight games and improved to 34-20 against [45 45 NL East 45] opponents . In [46 46 Houston 46] , [47 47 Tony Eusebio 47] 's eighth-inning sacrifice fly capped a comeback from a five-run deficit that gave the [48 48 Houston Astros 48] a 6-5 victory over the [49 49 Chicago Cubs 49] . The [50 50 Astros 50] trailed 5-0 after three innings , but scored three runs in the fourth and one in the sixth before taking the lead in the eighth . In [51 51 St Louis 51] , [52 52 Gary Sheffield 52] and [53 53 Devon White 53] each drove in two runs and [54 54 Mark Hutton 54] scattered four

hits over six innings to lead the [55 55 Florida Marlins 55] past the [56 56 St. Louis Cardinals 56] , 6-3 , [57 57 Sheffieled 57] , who was benched Monday , delivered a double down the left-field line in the first , scoring Luis Castillo and [58 58 Alex Arias 58] to put the [59 59 Marlins 59] ahead to stay . At [60 60 Colorado 60] , [61 61 Hal Morris 61] and [62 62 Eric Davis 62] each homered and [63 63 John Smiley 63] scattered six hits over 6 2/3 innings as the [64 64 Cincinnati Reds 64] defeated the [65 65 Colorado Rockies 65] 4-3 , snapping a four-game losing streak . The [66 66 Reds 66] took a one-run lead in the second inning when [67 67 Morris 67] led off with his 10th homer off starter [68 68 Armando Reynoso 68] (8-9) . They increased their bulge to 4-0 in the third when [69 69 Barry Larkin 69] drew a one-out walk , [70 70 Kevin Mitchell 70] singled and Davis launched his 22nd homer over the right-field wall .

1: 18213 Los_Angeles_Dodgers
2: 7954681 Montreal
3: 228180 Hideo_Nomo
4: 4516364 Greg_Gagne_(baseball)
5: 18213 Los_Angeles_Dodgers
6: 20153 Montreal_Expos
7: 18213 Los_Angeles_Dodgers
8: 20153 Montreal_Expos
9: 228180 Hideo_Nomo
10: 49728 San_Francisco
11: 4417184 Mike_Williams_(baseball)
12: 102264 Benito_Santiago
13: 23741 Philadelphia_Phillies
14: 28416 San_Francisco_Giants
15: 4417184 Mike_Williams_(baseball)
16: 4417184 Mike_Williams_(baseball)
17: 28416 San_Francisco_Giants
18: 25101 Pittsburgh
19: 3463887 Al_Martin
20: 4212649 Denny_Neagle
21: 391489 John_Smoltz
22: 24270 Pittsburgh_Pirates
23: 2140 Atlanta_Braves
24: 2140 Atlanta_Braves
25: 24270 Pittsburgh_Pirates
26: 391489 John_Smoltz
27: 4212649 Denny_Neagle
28: 2140 Atlanta_Braves
29: 27687 St._Louis
30: 1038021 Gary_Sheffield
31: 2536720 Devon_White_(baseball)
32: 3526509 Mark_Hutton
33: 11278 Miami_Marlins
34: 23408844 St._Louis_Cardinals
35: 7999625 Mark_Petkovsek
36: 11278 Miami_Marlins

```
37: 645042 New_York_City
38: 1035796 Steve_Finley
39: 28417 San_Diego_Padres
40: 1105870 Bobby_Valentine
41: 21728 New_York_Mets
42: 1046058 Wally_Joyner
43: 28417 San_Diego_Padres
44: 28417 San_Diego_Padres
45: 2456294 National_League_East
46: 13774 Houston
47: 8773129 Tony_Eusebio
48: 13894 Houston_Astros
49: 6654 Chicago_Cubs
50: 13894 Houston_Astros
51: 27687 St._Louis
52: 1038021 Gary_Sheffield
53: 2536720 Devon_White_(baseball)
54: 3526509 Mark_Hutton
55: 11278 Miami_Marlins
56: 23408844 St._Louis_Cardinals
57: 1038021 Gary_Sheffield
58: 9689206 Alex_Arias
59: 11278 Miami_Marlins
60: 5399 Colorado
61: 3594202 Hal_Morris
62: 1766916 Eric_Davis_(baseball)
63: 4359711 John_Smiley_(baseball)
64: 6671 Cincinnati_Reds
65: 6669 Colorado_Rockies
66: 6671 Cincinnati_Reds
67: 3594202 Hal_Morris
68: 6203994 Armando_Reynoso
69: 605895 Barry_Larkin
70: 643395 Kevin_Mitchell_(baseball)
1: 21728 New_York_Mets
    [winner: 4848143 New_York_Yankees, pem_pos=11,
final_scores=0.4419332, final_local_scores=0.37605757,
pem=-5.0692368, similarity=42.87932]
    gt: 21728 New_York_Mets, pem_pos=16,
final_scores=0.1298215, final_local_scores=0.032695055,
pem=-5.1311827, similarity=32.969566 ]
1: 18213 Los_Angeles_Dodgers
    [pem_pos=0, final_scores=0.5798189,
final_local_scores=0.52198434, pem=-0.20461318,
similarity=20.779486]
2: 7954681 Montreal
    [pem_pos=0, final_scores=0.56480825,
final_local_scores=0.67530847, pem=-0.11023275,
similarity=24.841728]
3: 228180 Hideo_Nomo
    [pem_pos=0, final_scores=0.6479275,
final_local_scores=0.66082644, pem=0.0, similarity=23.810093]
```

```
4: 4516364 Greg_Gagne_(baseball)
   [pem_pos=1, final_scores=0.5328801,
final_local_scores=0.56102407, pem=-0.19170615,
similarity=21.87432]
5: 18213 Los_Angeles_Dodgers
   [pem_pos=0, final_scores=0.56360596,
final_local_scores=0.5042742, pem=-0.105150424,
similarity=19.710085]
6: 20153 Montreal_Expos
   [pem_pos=0, final_scores=0.33179253,
final_local_scores=0.2605723, pem=-0.14979224,
similarity=12.680304]
7: 18213 Los_Angeles_Dodgers
   [pem_pos=0, final_scores=0.48318058,
final_local_scores=0.4164217, pem=-0.20461318,
similarity=17.629286]
8: 20153 Montreal_Expos
   [pem_pos=0, final_scores=0.4126264,
final_local_scores=0.34887105, pem=-0.034808703,
similarity=14.690016]
9: 228180 Hideo_Nomo
   [pem_pos=0, final_scores=0.5545617,
final_local_scores=0.56129086, pem=0.0, similarity=20.839752]
10: 49728 San_Francisco
    [pem_pos=0, final_scores=0.545932,
final_local_scores=0.61611456, pem=-0.09025385,
similarity=22.966616]
11: 4417184 Mike_Williams_(baseball)
    [pem_pos=4, final_scores=0.5525307,
final_local_scores=0.60145015, pem=-2.1541092,
similarity=33.752567]
12: 102264 Benito_Santiago
    [pem_pos=0, final_scores=0.5029163,
final_local_scores=0.50921464, pem=0.0, similarity=19.285694]
13: 23741 Philadelphia_Phillies
    [pem_pos=0, final_scores=0.512207,
final_local_scores=0.45113313, pem=-0.09993234,
similarity=18.095873]
14: 28416 San_Francisco_Giants
    [pem_pos=0, final_scores=0.6103587,
final_local_scores=0.5573331, pem=-0.07570464,
similarity=21.133339]
15: 4417184 Mike_Williams_(baseball)
    [pem_pos=5, final_scores=0.7679146,
final_local_scores=0.83672434, pem=-2.1541092,
similarity=40.773617]
16: 4417184 Mike_Williams_(baseball)
    [pem_pos=5, final_scores=0.75557905,
final_local_scores=0.82324964, pem=-2.1541092,
similarity=40.371506]
17: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.6159494,
```

```
final_local_scores=0.56344, pem=-4.409061,
similarity=44.881046]
18: 25101 Pittsburgh
    [pem_pos=0, final_scores=0.5144381,
final_local_scores=0.5785134, pem=-0.2986719,
similarity=22.97793]
19: 3463887 Al_Martin
    [pem_pos=0, final_scores=0.4451509,
final_local_scores=0.4966753, pem=0.0, similarity=18.911495]
20: 4212649 Denny_Neagle
    [pem_pos=0, final_scores=0.6723706,
final_local_scores=0.70977324, pem=0.0, similarity=25.270761]
21: 391489 John_Smoltz
    [pem_pos=0, final_scores=0.69707525,
final_local_scores=0.70329094, pem=0.0, similarity=25.077318]
22: 24270 Pittsburgh_Pirates
    [pem_pos=0, final_scores=0.4896301,
final_local_scores=0.42590332, pem=-0.15187988,
similarity=17.625463]
23: 2140 Atlanta_Braves
    [pem_pos=0, final_scores=0.4797318,
final_local_scores=0.41562587, pem=-0.06390974,
similarity=16.84037]
24: 2140 Atlanta_Braves
    [pem_pos=0, final_scores=0.5130957,
final_local_scores=0.45207083, pem=-0.28448334,
similarity=19.127472]
25: 24270 Pittsburgh_Pirates
    [pem_pos=1, final_scores=0.56044525,
final_local_scores=0.50325817, pem=-1.4743205,
similarity=27.125526]
26: 391489 John_Smoltz
    [pem_pos=0, final_scores=0.44440085,
final_local_scores=0.42728257, pem=0.0, similarity=16.840677]
27: 4212649 Denny_Neagle
    [pem_pos=0, final_scores=0.81693727,
final_local_scores=0.8707114, pem=0.0, similarity=30.073477]
28: 2140 Atlanta_Braves
    [pem_pos=0, final_scores=0.58987904,
final_local_scores=0.535945, pem=-0.28448334,
similarity=21.630445]
29: 27687 St._Louis
    [pem_pos=0, final_scores=0.42545122,
final_local_scores=0.45834833, pem=-0.092236795,
similarity=18.269339]
30: 1038021 Gary_Sheffield
    [pem_pos=0, final_scores=0.6932772,
final_local_scores=0.69232124, pem=0.0, similarity=24.74996]
31: 2536720 Devon_White_(baseball)
    [pem_pos=0, final_scores=0.62718695,
final_local_scores=0.66541785, pem=0.0, similarity=23.94711]
32: 3526509 Mark_Hutton
```

```
[pem_pos=0, final_scores=0.54993963,
final_local_scores=0.63321453, pem=0.0, similarity=22.986097]
33: 11278 Miami_Marlins
    [pem_pos=0, final_scores=0.46532738,
final_local_scores=0.40296513, pem=0.0, similarity=16.114996]
34: 23408844 St._Louis_Cardinals
    [pem_pos=0, final_scores=0.4388067,
final_local_scores=0.37437892, pem=-0.2366796,
similarity=16.549026]
35: 7999625 Mark_Petkovsek
    [pem_pos=0, final_scores=0.55950576,
final_local_scores=0.61955696, pem=0.0, similarity=22.57853]
36: 11278 Miami_Marlins
    [pem_pos=0, final_scores=0.3876763,
final_local_scores=0.3181431, pem=-0.22471714,
similarity=14.805784]
37: 645042 New_York_City
    [pem_pos=1, final_scores=0.40291744,
final_local_scores=0.47336906, pem=-1.7906687,
similarity=27.953924]
38: 1035796 Steve_Finley
    [pem_pos=0, final_scores=0.68269,
final_local_scores=0.68949956, pem=0.0, similarity=24.665754]
39: 28417 San_Diego_Padres
    [pem_pos=0, final_scores=0.5712947,
final_local_scores=0.51563746, pem=-0.05076447,
similarity=19.753428]
40: 1105870 Bobby_Valentine
    [pem_pos=0, final_scores=0.47766215,
final_local_scores=0.48134255, pem=0.0, similarity=18.453936]
41: 21728 New_York_Mets
    [pem_pos=0, final_scores=0.49190181,
final_local_scores=0.42789912, pem=-0.26481813,
similarity=18.299198]
42: 1046058 Wally_Joyner
    [pem_pos=0, final_scores=0.5564055,
final_local_scores=0.56935906, pem=0.0, similarity=21.080523]
43: 28417 San_Diego_Padres
    [pem_pos=0, final_scores=0.52254933,
final_local_scores=0.46239054, pem=-0.049586933,
similarity=18.15803]
44: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.61311525,
final_local_scores=0.56132007, pem=-3.9307637,
similarity=42.21673]
45: 2456294 National_League_East
    [pem_pos=0, final_scores=0.41756368,
final_local_scores=0.4167241, pem=0.0, similarity=16.52559]
46: 13774 Houston
    [pem_pos=0, final_scores=0.64693946,
final_local_scores=0.71318436, pem=-0.3232624,
similarity=27.130508]
```

47: 8773129 Tony_Eusebio
[pem_pos=0, final_scores=0.5409147,
final_local_scores=0.58028066, pem=0.0, similarity=21.406445]
48: 13894 Houston_Astros
[pem_pos=0, final_scores=0.41530782,
final_local_scores=0.34311247, pem=-0.049621258,
similarity=14.5987215]
49: 6654 Chicago_Cubs
[pem_pos=0, final_scores=0.4279613,
final_local_scores=0.35936236, pem=-0.085107334,
similarity=15.27663]
50: 13894 Houston_Astros
[pem_pos=0, final_scores=0.50477916,
final_local_scores=0.44084626, pem=-0.294573,
similarity=18.84738]
51: 27687 St._Louis
[pem_pos=0, final_scores=0.4385695,
final_local_scores=0.47267807, pem=-0.092236795,
similarity=18.696966]
52: 1038021 Gary_Sheffield
[pem_pos=0, final_scores=0.69709533,
final_local_scores=0.696492, pem=0.0, similarity=24.874424]
53: 2536720 Devon_White_(baseball)
[pem_pos=0, final_scores=0.6234502,
final_local_scores=0.661336, pem=0.0, similarity=23.825298]
54: 3526509 Mark_Hutton
[pem_pos=0, final_scores=0.5562925,
final_local_scores=0.64015406, pem=0.0, similarity=23.193188]
55: 11278 Miami_Marlins
[pem_pos=0, final_scores=0.48273873,
final_local_scores=0.42198437, pem=0.0, similarity=16.682568]
56: 23408844 St._Louis_Cardinals
[pem_pos=0, final_scores=0.5398342,
final_local_scores=0.48473614, pem=-0.2366796,
similarity=19.842306]
57: 1038021 Gary_Sheffield
[pem_pos=0, final_scores=1.4656866,
final_local_scores=1.5380199, pem=0.0, similarity=49.987293]
58: 9689206 Alex_Arias
[pem_pos=0, final_scores=0.43736917,
final_local_scores=0.47507268, pem=0.0, similarity=18.266829]
59: 11278 Miami_Marlins
[pem_pos=0, final_scores=0.39594358,
final_local_scores=0.32717386, pem=-0.22471714,
similarity=15.07528]
60: 5399 Colorado
[pem_pos=0, final_scores=0.4325996,
final_local_scores=0.53677815, pem=-0.26926568,
similarity=21.572554]
61: 3594202 Hal_Morris
[pem_pos=0, final_scores=0.43894476,
final_local_scores=0.4760188, pem=0.0, similarity=18.295063]

```
62: 1766916 Eric_Davis_(baseball)
    [pem_pos=0, final_scores=0.624368,
final_local_scores=0.65425944, pem=-0.1638126,
similarity=24.50496]
63: 4359711 John_Smiley_(baseball)
    [pem_pos=1, final_scores=0.671443,
final_local_scores=0.7514152, pem=-0.37469345,
similarity=28.551083]
64: 6671 Cincinnati_Reds
    [pem_pos=0, final_scores=0.60386246,
final_local_scores=0.5543471, pem=-0.106723964,
similarity=21.212917]
65: 6669 Colorado_Rockies
    [pem_pos=0, final_scores=0.515432,
final_local_scores=0.45830423, pem=-0.10158596,
similarity=18.318867]
66: 6671 Cincinnati_Reds
    [pem_pos=1, final_scores=0.48276734,
final_local_scores=0.42206907, pem=-1.0258155,
similarity=22.26364]
67: 3594202 Hal_Morris
    [pem_pos=0, final_scores=0.67279327,
final_local_scores=0.7314627, pem=0.0, similarity=25.91802]
68: 6203994 Armando_Reynoso
    [pem_pos=0, final_scores=0.4863106,
final_local_scores=0.54840314, pem=0.0, similarity=20.455156]
69: 605895 Barry_Larkin
    [pem_pos=0, final_scores=0.6915063,
final_local_scores=0.70562124, pem=0.0, similarity=25.146858]
70: 643395 Kevin_Mitchell_(baseball)
    [pem_pos=0, final_scores=0.48707885,
final_local_scores=0.5006439, pem=-0.3107999,
similarity=20.720104]
1: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.4419332,
final_local_scores=0.37605757, pem=-5.0692368,
similarity=42.87932]
```

Sample: b'653_RUGBY'

[1 1 RUGBY UNION 1] - [2 2 NEW ZEALAND 2] RECALL [3 3 MEHRTE
NS 3] FOR FINAL TEST . [4 4 JOHANNESBURG 4] 1996-08-28 The [5
5 New Zealand 5] rugby selectors recalled fly-half [6 6 Andrew
Mehrtens 6] on Wednesday when they announced their team for
the third and final test in [7 7 Johannesburg 7] on
Saturday . He returns in place of [8 8 Simon Culhane 8] who
broke a wrist in the [9 9 All Blacks 9] ' series-clinching
victory in [10 10 Pretoria 10] on Saturday . [11 11 Mehrtens 1
1] played in the last TriNations test in [12 12 Cape Town 12]
but missed the first two tests in the current series after
tearing a cartilage in his knee while training , an injury
which needed a small operation . Lock [13 13 Ian Jones 13]
and wing [14 14 Jeff Wilson 14] have also been named in the

team despite doubts over their fitness . Jones has a knee injury while [15 15 Wilson 15] is suffering from a viral infection . Blair Larsen or the uncapped Glenn Taylor are on standby to replace Jones and , with [16 16 Jonah Lomu 16] out of action with a shoulder injury picked up in Tuesday 's drawn match against [17 17 Griqualand West 17] , [18 18 Eric Rush 18] is favourite to play should [19 19 Wilson 19] fail to recover . Team : 15 - [20 20 Christian Cullen 20] , 14 - [21 21 Jeff Wilson 21] , 13 - [22 22 Walter Little 22] , 12 - [23 23 Frank Bunce 23] , 11 - [24 24 Glen Osborne 24] ; 10 - [25 25 Andrew Mehrtens 25] , 9 - [26 26 Justin Marshall 26] ; 8 - [27 27 Zinzan Brooke 27] , 7 - [28 28 Josh Kronfeld 28] , 6 - [29 29 Michael Jones 29] , 5 - [30 30 Ian Jones 30] , 4 - [31 31 Robin Brooke 31] , 3 - [32 32 Olo Brown 32] , 2 - [33 33 Sean Fitzpatrick 33] (captain) , 1 - [34 34 Craig Dowd 34] .

1: 25405 Rugby_union
2: 154957 New_Zealand_national_rugby_union_team
3: 860422 Andrew_Mehrtens
4: 58893 Johannesburg
5: 154957 New_Zealand_national_rugby_union_team
6: 860422 Andrew_Mehrtens
7: 58893 Johannesburg
8: 5210833 Simon_Culhane
9: 154957 New_Zealand_national_rugby_union_team
10: 25002 Pretoria
11: 860422 Andrew_Mehrtens
12: 6653 Cape_Town
13: 5974513 Ian_Jones_(rugby_union)
14: 905152 Jeff_Wilson_(sportsman)
15: 905152 Jeff_Wilson_(sportsman)
16: 264230 Jonah_Lomu
17: 888197 Griqualand_West
18: 3439181 Eric_Rush
19: 905152 Jeff_Wilson_(sportsman)
20: 346051 Christian_Cullen
21: 905152 Jeff_Wilson_(sportsman)
22: 27620822 Walter_Little_(rugby_player)
23: 6093538 Frank_Bunce
24: 11940740 Glen_Osborne
25: 860422 Andrew_Mehrtens
26: 1809625 Justin_Marshall
27: 740724 Zinzan_Brooke
28: 1309540 Josh_Kronfeld
29: 1294466 Michael_Jones_(rugby_union)
30: 5974513 Ian_Jones_(rugby_union)
31: 9033319 Robin_Brooke
32: 5821690 Olo_Brown
33: 1444074 Sean_Fitzpatrick
34: 15821736 Craig_Dowd
1: 25405 Rugby_union
[pem_pos=0, final_scores=0.52496773, final_local_scores=0.5353912, pem=0.0, similarity=20.066854]

```
2: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=4, final_scores=0.62644154,
final_local_scores=0.59099257, pem=-3.8192244,
similarity=42.495644]
3: 860422 Andrew_Mehrtens
    [pem_pos=0, final_scores=0.7443252,
final_local_scores=0.74886066, pem=0.0, similarity=26.437208]
4: 58893 Johannesburg
    [pem_pos=0, final_scores=0.47741365,
final_local_scores=0.52418137, pem=-0.09562059,
similarity=20.25233]
5: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=4, final_scores=0.6051231,
final_local_scores=0.5677054, pem=-3.8192244,
similarity=41.80071]
6: 860422 Andrew_Mehrtens
    [pem_pos=0, final_scores=0.740848,
final_local_scores=0.74506223, pem=0.0, similarity=26.323856]
7: 58893 Johannesburg
    [pem_pos=0, final_scores=0.5513337,
final_local_scores=0.6049278, pem=-0.09562059,
similarity=22.661964]
8: 5210833 Simon_Culhane
    [pem_pos=0, final_scores=0.6745334,
final_local_scores=0.73427546, pem=0.0, similarity=26.001957]
9: 154957 New_Zealand_national_rugby_union_team
    [pem_pos=0, final_scores=0.5266767,
final_local_scores=0.48201466, pem=0.0, similarity=18.473991]
10: 25002 Pretoria
    [pem_pos=0, final_scores=0.44020802,
final_local_scores=0.48030734, pem=0.0, similarity=18.423042]
11: 860422 Andrew_Mehrtens
    [pem_pos=0, final_scores=0.6919852,
final_local_scores=0.69168705, pem=0.0, similarity=24.731033]
12: 6653 Cape_Town
    [pem_pos=0, final_scores=0.43631977,
final_local_scores=0.48526186, pem=0.0, similarity=18.570894]
13: 5974513 Ian_Jones_(rugby_union)
    [pem_pos=0, final_scores=0.42501253,
final_local_scores=0.45257443, pem=-0.0020345869,
similarity=17.6065]
14: 905152 Jeff_Wilson_(sportsman)
    [pem_pos=0, final_scores=0.75580186,
final_local_scores=0.7641804, pem=0.0, similarity=26.89438]
15: 905152 Jeff_Wilson_(sportsman)
    [pem_pos=0, final_scores=0.59194237,
final_local_scores=0.5851888, pem=0.0, similarity=21.552914]
16: 264230 Jonah_Lomu
    [pem_pos=0, final_scores=0.5495011,
final_local_scores=0.53265786, pem=0.0, similarity=19.985287]
17: 888197 Griqualand_West
    [pem_pos=0, final_scores=0.38488668,
```

```
final_local_scores=0.4609335, pem=-0.045276318,
similarity=18.091108]
18: 3439181 Eric_Rush
    [pem_pos=0, final_scores=0.6210561,
final_local_scores=0.68196505, pem=0.0, similarity=24.44091]
19: 905152 Jeff_Wilson_(sportsman)
    [pem_pos=0, final_scores=0.66747886,
final_local_scores=0.66770095, pem=0.0, similarity=24.015242]
20: 346051 Christian_Cullen
    [pem_pos=0, final_scores=0.6699949,
final_local_scores=0.6784237, pem=0.0, similarity=24.33523]
21: 905152 Jeff_Wilson_(sportsman)
    [pem_pos=0, final_scores=0.68559545,
final_local_scores=0.6874906, pem=0.0, similarity=24.605803]
22: 27620822 Walter_Little_(rugby_player)
    [pem_pos=0, final_scores=0.42254704,
final_local_scores=0.46388787, pem=0.0, similarity=17.933052]
23: 6093538 Frank_Bunce
    [pem_pos=0, final_scores=0.7620496,
final_local_scores=0.7975007, pem=0.0, similarity=27.888725]
24: 11940740 Glen_Osborne
    [pem_pos=1, final_scores=0.5650555,
final_local_scores=0.65411806, pem=-0.27698678,
similarity=25.116198]
25: 860422 Andrew_Mehrtens
    [pem_pos=0, final_scores=0.6936968,
final_local_scores=0.6935568, pem=0.0, similarity=24.78683]
26: 1809625 Justin_Marshall
    [pem_pos=0, final_scores=0.74116737,
final_local_scores=0.7812789, pem=0.0, similarity=27.404634]
27: 740724 Zinzan_Brooke
    [pem_pos=0, final_scores=0.7246375,
final_local_scores=0.7664344, pem=0.0, similarity=26.961643]
28: 1309540 Josh_Kronfeld
    [pem_pos=0, final_scores=0.51196635,
final_local_scores=0.52848744, pem=0.0, similarity=19.860832]
29: 1294466 Michael_Jones_(rugby_union)
    [pem_pos=0, final_scores=0.3637553,
final_local_scores=0.3601169, pem=-1.0030928,
similarity=20.291294]
30: 5974513 Ian_Jones_(rugby_union)
    [pem_pos=0, final_scores=0.58464307,
final_local_scores=0.6269466, pem=-0.0020345869,
similarity=22.810114]
31: 9033319 Robin_Brooke
    [pem_pos=0, final_scores=0.5616339,
final_local_scores=0.6369613, pem=0.0, similarity=23.097908]
32: 5821690 Olo_Brown
    [pem_pos=0, final_scores=0.6005773,
final_local_scores=0.7071757, pem=0.0, similarity=25.193245]
33: 1444074 Sean_Fitzpatrick
    [pem_pos=0, final_scores=0.65605205,
```

```
final_local_scores=0.6816919, pem=0.0, similarity=24.432758]
34: 15821736 Craig_Dowd
    [pem_pos=0, final_scores=0.5460929,
final_local_scores=0.6041924, pem=0.0, similarity=22.120018]

Sample: b'654_HOCKEY'
HOCKEY - BONNET TAKES OVER AS SOUTH AFRICAN COACH . [1 1 JOH
ANNESBURG 1] 1996-08-28 Former South African captain Giles
Bonnet was named by the South African Hockey Association on
Wednesday as the new coach of the men 's national side .
Bonnet , who has been coaching the [2 2 KwaZulu-Natal 2]
provincial team , takes over from [3 3 Englishman 3] Gavin
Featherstone who took 1 South Africa to 10th place in the 201
ympic Games in [4 4 Atlanta 4] . Featherstone , a former
Britain captain , has accepted a coaching position with a
women 's team in [5 5 Ireland 5] .
1: 58893 Johannesburg
2: 193596 KwaZulu-Natal
3: 9316 England
4: 3138 Atlanta
5: 14560 Republic_of_Ireland
1: 18523612 South_Africa_national_field_hockey_team
    [winner: 376115 South_Africa_national_rugby_union_team,
pem_pos=3, final_scores=-0.12948257,
final_local_scores=-0.126791, pem=-3.598493,
similarity=19.875181|
    RECALL MISS: 18523612
South_Africa_national_field_hockey_team]
2: 72311 1996_Summer_Olympics
    [winner: 22576 Olympic_Games, pem_pos=0,
final_scores=-0.25645953, final_local_scores=-0.28834814,
pem=-0.5587218, similarity=-1.4767549|
    RECALL MISS: 72311 1996_Summer_Olympics]
1: 58893 Johannesburg
    [pem_pos=0, final_scores=0.51023835,
final_local_scores=0.50036913, pem=-0.09562059,
similarity=19.541727]
2: 193596 KwaZulu-Natal
    [pem_pos=0, final_scores=0.2723797,
final_local_scores=0.2986141, pem=0.0, similarity=13.000954]
3: 9316 England
    [pem_pos=1, final_scores=0.21988098,
final_local_scores=0.19379926, pem=-0.8308641,
similarity=14.391438]
4: 3138 Atlanta
    [pem_pos=0, final_scores=0.44033772,
final_local_scores=0.46433944, pem=-0.23152597,
similarity=19.205603]
5: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.4306891,
final_local_scores=0.45141774, pem=-1.5040284,
similarity=25.740059]
```

Sample: b'655_CRICKET'

CRICKET - [1 1 GIBBS 1] GETS INTERNATIONAL CALL UP . [2 2 JO HANNESBURG 2] 1996-08-28 [3 3 Western Province 3] batsman [4 4 Herschelle Gibbs 4] was the only uncapped player in [5 5 South Africa 5] 's 14-man squad named on Wednesday for a quadrangular one-day series in [6 6 Kenya 6] next month . [7 7 Kenya 7] , [8 8 South Africa 8] , [9 9 Pakistan 9] and [10 10 Sri Lanka 10] will take part in the series . National coach [1 1 11 Bob Woolmer 11] said [12 12 Gibbs 12] , 22 , had been rewarded for a tremendous tour of [13 13 England 13] with the [14 14 South African 14] A team earlier this year . " I 've known [15 15 Herschelle 15] since he was 11 years old and he showed in [16 16 England 16] how he has matured . His 170 against the [17 17 MCC 17] was an innings of supreme class against the best bowling attack we faced all tour , " [18 18 W oolmer 18] told a news conference . . " We were not able to consider [19 19 Jacques Kallis 19] , [20 20 Paul Adams 20] and [21 21 Shaun Pollock 21] due to injury and the replacements have all come from the A tour and it 's great that they are all in form . " Spin-bowling all-rounders [22 22 Nicky Boje 22] and [23 23 Derek Crookes 23] replace [24 24 Po llock 24] and Adams , while [25 25 Gibbs 25] comes in for his [26 26 Western Province 26] colleague [27 27 Kallis 27] . Squad : [28 28 Hansie Cronje 28] (captain) , [29 29 Craig Ma tthews 29] (vice-captain) , [30 30 Dave Richardson 30] , [31 31 Brian McMillan 31] , [32 32 Gary Kirsten 32] , [33 33 Andr ew Hudson 33] , [34 34 Pat Symcox 34] , [35 35 Jonty Rhodes 35] , [36 36 Allan Donald 36] , [37 37 Fanie de Villiers 37] , [38 38 Daryll Cullinan 38] , [39 39 Derek Crookes 39] , Herschelle Gibbs , [40 40 Nicky Boje 40] .

1: 897189 Herschelle_Gibbs

2: 58893 Johannesburg

3: 6904363 Western_Province_cricket_team_(South_Africa)

4: 897189 Herschelle_Gibbs

5: 759263 South_Africa_national_cricket_team

6: 188171 Kenya

7: 1051029 Kenya_national_cricket_team

8: 759263 South_Africa_national_cricket_team

9: 759295 Pakistan_national_cricket_team

10: 397246 Sri_Lanka_national_cricket_team

11: 1485764 Bob_Woolmer

12: 897189 Herschelle_Gibbs

13: 9316 England

14: 17416221 South_Africa

15: 897189 Herschelle_Gibbs

16: 9316 England

17: 307442 Marylebone_Cricket_Club

18: 1485764 Bob_Woolmer

19: 571857 Jacques_Kallis

20: 1677731 Paul_Adams_(cricketer)

21: 897244 Shaun_Pollock

```
22: 3060990 Nicky_Boje
23: 4305466 Derek_Crookes
24: 897244 Shaun_Pollock
25: 897189 Herschelle_Gibbs
26: 6904363 Western_Province_cricket_team_(South_Africa)
27: 571857 Jacques_Kallis
28: 13961 Hansie_Cronje
29: 3831183 Craig_Matthews
30: 3182138 Dave_Richardson
31: 3635821 Brian_McMillan
32: 897261 Gary_Kirsten
33: 3026499 Andrew_Hudson
34: 3060256 Pat_Symcox
35: 4287169 Jonty_Rhodes
36: 629053 Allan_Donald
37: 3174359 Fanie_de_Villiers
38: 3490162 Daryll_Cullinan
39: 4305466 Derek_Crookes
40: 3060990 Nicky_Boje
1: 897189 Herschelle_Gibbs
    [pem_pos=0, final_scores=1.2278273,
final_local_scores=1.2631478, pem=0.0, similarity=41.784565]
2: 58893 Johannesburg
    [pem_pos=0, final_scores=0.48873365,
final_local_scores=0.54105026, pem=-0.09562059,
similarity=20.755732]
3: 6904363 Western_Province_cricket_team_(South_Africa)
    [pem_pos=1, final_scores=0.38106567,
final_local_scores=0.36191744, pem=-1.3355793,
similarity=22.15314]
4: 897189 Herschelle_Gibbs
    [pem_pos=0, final_scores=1.0170512,
final_local_scores=1.03021, pem=0.0, similarity=34.833237]
5: 759263 South_Africa_national_cricket_team
    [pem_pos=1, final_scores=0.6788357,
final_local_scores=0.65681726, pem=-3.0105038,
similarity=40.06204]
6: 188171 Kenya
    [pem_pos=0, final_scores=0.43446642,
final_local_scores=0.54279464, pem=-0.27389407,
similarity=21.777266]
7: 1051029 Kenya_national_cricket_team
    [pem_pos=1, final_scores=0.52382356,
final_local_scores=0.5206106, pem=-2.794723,
similarity=34.82391]
8: 759263 South_Africa_national_cricket_team
    [pem_pos=1, final_scores=0.73966306,
final_local_scores=0.7232619, pem=-3.0105038,
similarity=42.044876]
9: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7783478,
final_local_scores=0.7840156, pem=-2.7556365,
```

```
similarity=42.47188]
10: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.8505115,
final_local_scores=0.8646882, pem=-2.2667584,
similarity=42.22072]
11: 1485764 Bob_Woolmer
    [pem_pos=0, final_scores=0.6277669,
final_local_scores=0.6411755, pem=0.0, similarity=23.22367]
12: 897189 Herschelle_Gibbs
    [pem_pos=0, final_scores=1.4129632,
final_local_scores=1.4653807, pem=0.0, similarity=47.819595]
13: 9316 England
    [pem_pos=0, final_scores=0.51756173,
final_local_scores=0.61611044, pem=-0.19053856,
similarity=23.511856]
14: 17416221 South_Africa
    [pem_pos=0, final_scores=0.4495231,
final_local_scores=0.51531637, pem=-0.1389451,
similarity=20.223385]
15: 897189 Herschelle_Gibbs
    [pem_pos=0, final_scores=0.69264174,
final_local_scores=0.67584187, pem=0.0, similarity=24.258183]
16: 9316 England
    [pem_pos=0, final_scores=0.40986407,
final_local_scores=0.4984671, pem=-0.19053856,
similarity=20.001143]
17: 307442 Marylebone_Cricket_Club
    [pem_pos=0, final_scores=0.66299623,
final_local_scores=0.68811727, pem=-0.18132487,
similarity=25.610579]
18: 1485764 Bob_Woolmer
    [pem_pos=0, final_scores=0.5637143,
final_local_scores=0.57570857, pem=0.0, similarity=21.270004]
19: 571857 Jacques_Kallis
    [pem_pos=0, final_scores=0.7842545,
final_local_scores=0.78477067, pem=0.0, similarity=27.508835]
20: 1677731 Paul_Adams_(cricketer)
    [pem_pos=1, final_scores=0.6075194,
final_local_scores=0.66244966, pem=-0.6008317,
similarity=27.12595]
21: 897244 Shaun_Pollock
    [pem_pos=0, final_scores=1.2278781,
final_local_scores=1.265748, pem=0.0, similarity=41.862156]
22: 3060990 Nicky_Boje
    [pem_pos=0, final_scores=0.69963026,
final_local_scores=0.7251678, pem=0.0, similarity=25.730167]
23: 4305466 Derek_Crookes
    [pem_pos=0, final_scores=0.49259818,
final_local_scores=0.5676315, pem=0.0, similarity=21.028969]
24: 897244 Shaun_Pollock
    [pem_pos=0, final_scores=1.4976647,
final_local_scores=1.5634208, pem=0.0, similarity=50.745304]
```

```
25: 897189 Herschelle_Gibbs
    [pem_pos=0, final_scores=1.5220261,
final_local_scores=1.5845153, pem=0.0, similarity=51.37481]
26: 6904363 Western_Province_cricket_team_(South_Africa)
    [pem_pos=1, final_scores=0.46695065,
final_local_scores=0.45573378, pem=-1.3355793,
similarity=24.952805]
27: 571857 Jacques_Kallis
    [pem_pos=0, final_scores=0.76675004,
final_local_scores=0.7656497, pem=0.0, similarity=26.938227]
28: 13961 Hansie_Cronje
    [pem_pos=0, final_scores=0.6870644,
final_local_scores=0.6921516, pem=0.0, similarity=24.744898]
29: 3831183 Craig_Matthews
    [pem_pos=0, final_scores=0.4429564,
final_local_scores=0.5320438, pem=0.0, similarity=19.96696]
30: 3182138 Dave_Richardson
    [pem_pos=0, final_scores=0.5762271,
final_local_scores=0.6394301, pem=0.0, similarity=23.171583]
31: 3635821 Brian_McMillan
    [pem_pos=0, final_scores=0.6347258,
final_local_scores=0.697312, pem=0.0, similarity=24.898893]
32: 897261 Gary_Kirsten
    [pem_pos=0, final_scores=0.9486926,
final_local_scores=0.9856644, pem=0.0, similarity=33.503906]
33: 3026499 Andrew_Hudson
    [pem_pos=0, final_scores=0.52322465,
final_local_scores=0.5708179, pem=0.0, similarity=21.124058]
34: 3060256 Pat_Symcox
    [pem_pos=0, final_scores=0.7046785,
final_local_scores=0.7731074, pem=0.0, similarity=27.16078]
35: 4287169 Jonty_Rhodes
    [pem_pos=0, final_scores=0.7254121,
final_local_scores=0.73742634, pem=0.0, similarity=26.095985]
36: 629053 Allan_Donald
    [pem_pos=0, final_scores=0.858491,
final_local_scores=0.8801059, pem=0.0, similarity=30.353828]
37: 3174359 Fanie_de_Villiers
    [pem_pos=0, final_scores=0.71604663,
final_local_scores=0.76261306, pem=0.0, similarity=26.847609]
38: 3490162 Daryll_Cullinan
    [pem_pos=0, final_scores=0.71897703,
final_local_scores=0.779862, pem=0.0, similarity=27.36235]
39: 4305466 Derek_Crookes
    [pem_pos=0, final_scores=0.49768227,
final_local_scores=0.5731851, pem=0.0, similarity=21.194698]
40: 3060990 Nicky_Boje
    [pem_pos=0, final_scores=0.6524661,
final_local_scores=0.6736481, pem=0.0, similarity=24.192717]
```

Sample: b'656_BASKETBALL'
BASKETBALL - [1 OLYMPIAKOS 1] BEAT [1 1 RED STAR 1] 71-57 . [

2 2 BELGRADE 2] 1996-08-28 Olympiakos of [3 3 Greece 3] beat [4 4 Yugoslavia 4] 's [5 5 Red Star 5] 71-57 (halftime 40-34) in the first match of an international club basketball tournament on Wednesday . [6 6 Partizan 6] ([2 1 Yugoslavia 2]) , Alba ([7 7 Germany 7]) , Dinamo ([8 8 Russia 8]) and [9 9 Benetton 9] ([10 10 Italy 10]) are also taking part in the event which continues until Saturday .

1: 5604103 KK_Crvena_zvezda
2: 55904 Belgrade
3: 12108 Greece
4: 34244 Yugoslavia
5: 5604103 KK_Crvena_zvezda
6: 5551651 KK_Partizan
7: 11867 Germany
8: 25391 Russia
9: 3450530 Treviso_Basket
10: 14532 Italy
1: 29265 Serbia
[winner: 34244 Yugoslavia, pem_pos=0,
final_scores=0.23095821, final_local_scores=0.17862645,
pem=-0.6536551, similarity=12.97496]
RECALL MISS: 29265 Serbia]

1: 5604103 KK_Crvena_zvezda
[pem_pos=5, final_scores=0.2309237,
final_local_scores=0.2245251, pem=-2.4627666,
similarity=24.182898]
2: 55904 Belgrade
[pem_pos=0, final_scores=0.4916765,
final_local_scores=0.46749783, pem=-0.05193398,
similarity=18.323206]
3: 12108 Greece
[pem_pos=0, final_scores=0.78963214,
final_local_scores=0.8241129, pem=-0.27607638,
similarity=30.18423]
4: 34244 Yugoslavia
[pem_pos=0, final_scores=0.38131237,
final_local_scores=0.3428655, pem=-0.6536551,
similarity=17.87618]
5: 5604103 KK_Crvena_zvezda
[pem_pos=5, final_scores=0.3078671,
final_local_scores=0.30098546, pem=-2.4627666,
similarity=26.464628]
6: 5551651 KK_Partizan
[pem_pos=1, final_scores=0.32331663,
final_local_scores=0.3226314, pem=-1.689159,
similarity=22.903585]
7: 11867 Germany
[pem_pos=0, final_scores=0.7392153,
final_local_scores=0.78462327, pem=-0.15529114,
similarity=28.348934]
8: 25391 Russia
[pem_pos=0, final_scores=0.7251647,

```
final_local_scores=0.7839818, pem=-0.20299491,
similarity=28.58921]
9: 3450530 Treviso_Basket
    [pem_pos=4, final_scores=0.43182045,
final_local_scores=0.5016735, pem=-2.1818364,
similarity=30.925817]
10: 14532 Italy
    [pem_pos=0, final_scores=0.5840551,
final_local_scores=0.61195284, pem=-0.20612453,
similarity=23.472544]
1: 1229550 Olympiacos_F.C.
    [pem_pos=0, final_scores=0.10141568,
final_local_scores=0.09260118, pem=-0.4098014,
similarity=9.08168]
2: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.23095821,
final_local_scores=0.17862645, pem=-0.6536551,
similarity=12.97496]
```

Sample: b'657_SOCCER'
SOCCER - [1 1 RUSSIA 1] AND [2 2 BRAZIL 2] DRAW 2-2 IN FRIENDLY . [3 3 MOSCOW 3] 1996-08-28 [4 4 Russia 4] and [5 5 Brazil 5] drew 2-2 (halftime 1-0) in a friendly soccer international on Wednesday . Scorers : [6 6 Russia 6] - [7 7 Yuri Nikiforov 7] (18th minute) , Vladislav Rodimov (80th) [8 8 Brazil 8] - Donizetti (47th) , [9 9 Ronaldo 9] (85th) Attendence : 20,000

```
1: 25391 Russia
2: 3383 Brazil
3: 19004 Moscow
4: 25127146 Russia_national_football_team
5: 149286 Brazil_national_football_team
6: 25127146 Russia_national_football_team
7: 6004170 Yuriy_Nikiforov
8: 149286 Brazil_national_football_team
9: 66164 Ronaldo
1: 25391 Russia
    [pem_pos=0, final_scores=0.4574619,
final_local_scores=0.49513727, pem=-0.20299491,
similarity=19.969515]
2: 3383 Brazil
    [pem_pos=0, final_scores=0.54078287,
final_local_scores=0.5963715, pem=-0.21313342,
similarity=23.04568]
3: 19004 Moscow
    [pem_pos=0, final_scores=0.6495163,
final_local_scores=0.7093606, pem=-0.106582455,
similarity=25.83806]
4: 25127146 Russia_national_football_team
    [pem_pos=2, final_scores=0.4533986,
final_local_scores=0.41851234, pem=-3.5373242,
similarity=35.815475]
```

```
5: 149286 Brazil_national_football_team
    [pem_pos=1, final_scores=0.5952569,
final_local_scores=0.572118, pem=-2.5928838,
similarity=35.26336]
6: 25127146 Russia_national_football_team
    [pem_pos=2, final_scores=0.56429404,
final_local_scores=0.5396488, pem=-3.5373242,
similarity=39.430428]
7: 6004170 Yuriy_Nikiforov
    [pem_pos=0, final_scores=0.580186,
final_local_scores=0.6601805, pem=0.0, similarity=23.790815]
8: 149286 Brazil_national_football_team
    [pem_pos=1, final_scores=0.6472878,
final_local_scores=0.62895375, pem=-2.5928838,
similarity=36.959454]
9: 66164 Ronaldo
    [pem_pos=0, final_scores=0.3214875,
final_local_scores=0.34595403, pem=-0.10865249,
similarity=15.00454]
```

Sample: b'658_SQUASH'

SQUASH - HONG OPEN FIRST ROUND RESULTS . [1 1 HONG KONG 1]
1996-08-28 First round results in the Hong Kong Open on
Wednesday (prefix denotes seeding) : 2 - [2 2 Rodney Eyles 2
] ([3 3 Australia 3]) beat Zarak Jahan Khan ([4 4 Pakistan
4]) 15-6 8-15 15-10 7-15 15-12 4 - [5 5 Peter Nicol 5] ([6 6
Scotland 6]) beat Julian Wellings ([7 7 England 7]) 15-8
15-7 15-6 [8 8 Derek Ryan 8] ([9 9 Ireland 9]) beat 5 - [10
10 Simon Parke 10] ([11 11 England 11]) 15-11 15-11 2-15
15-11 7 - [12 12 Chris Walker 12] ([13 13 England 13]) beat
Julien Bonetat ([14 14 France 14]) 15-12 15-6 15-2 [15 15 Jo
nathon Power 15] ([16 16 Canada 16]) beat [17 17 Ahmed Bara
da 17] ([18 18 Egypt 18]) 11-15 8-15 15-13 15-11 15-2 [19 19
Amr Shabana 19] ([20 20 Egypt 20]) beat [21 21 John White 21
] ([22 22 Australia 22]) 10-15 15-9 15-10 16-17 15-1 [23 23
Paul Johnson 23] ([24 24 England 24]) beat Tony Hands ([25
25 England 25]) 12-15 15-11 7-15 15-6 15-11 Zubair Jahan
Khan ([26 26 Pakistan 26]) beat Faheem Khan ([27 27 Hong Ko
ng 27]) 12-15 15-10 15-10 15-10 R

```
1: 13404 Hong_Kong
2: 3428311 Rodney_Eyles
3: 4689264 Australia
4: 23235 Pakistan
5: 2947276 Peter_Nicol
6: 26994 Scotland
7: 9316 England
8: 24449519 Derek_Ryan_(squash_player)
9: 14560 Republic_of_Ireland
10: 3451294 Simon_Parke
11: 9316 England
12: 3672106 Chris_Walker_(squash_player)
13: 9316 England
```

```
14: 5843419 France
15: 1529068 Jonathon_Power
16: 5042916 Canada
17: 3672222 Ahmed_Barada
18: 8087628 Egypt
19: 3451400 Amr_Shabana
20: 8087628 Egypt
21: 3671720 John_White_(squash_player)
22: 4689264 Australia
23: 7434625 Paul_Johnson_(squash_player)
24: 9316 England
25: 9316 England
26: 23235 Pakistan
27: 13404 Hong_Kong
1: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.50526863,
final_local_scores=0.49631143, pem=-0.22402541,
similarity=20.118921]
2: 3428311 Rodney_Eyles
    [pem_pos=0, final_scores=0.6586285,
final_local_scores=0.77631503, pem=0.0, similarity=27.256502]
3: 4689264 Australia
    [pem_pos=0, final_scores=0.6897735,
final_local_scores=0.6940528, pem=-0.17168477,
similarity=25.735283]
4: 23235 Pakistan
    [pem_pos=0, final_scores=0.6575188,
final_local_scores=0.69048184, pem=-0.21944217,
similarity=25.888428]
5: 2947276 Peter_Nicol
    [pem_pos=0, final_scores=0.6851236,
final_local_scores=0.78487647, pem=0.0, similarity=27.511992]
6: 26994 Scotland
    [pem_pos=0, final_scores=0.46751148,
final_local_scores=0.46356606, pem=-0.22905664,
similarity=19.169094]
7: 9316 England
    [pem_pos=0, final_scores=0.56232107,
final_local_scores=0.5361131, pem=-0.19053856,
similarity=21.124575]
8: 24449519 Derek_Ryan_(squash_player)
    [pem_pos=1, final_scores=0.49631542,
final_local_scores=0.62746435, pem=-1.3862944,
similarity=30.353386]
9: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.5187003,
final_local_scores=0.54097897, pem=-1.5040284,
similarity=28.412745]
10: 3451294 Simon_Parke
    [pem_pos=0, final_scores=0.611866,
final_local_scores=0.7559167, pem=0.0, similarity=26.647776]
11: 9316 England
```

```
[pem_pos=0, final_scores=0.54146296,
final_local_scores=0.5133288, pem=-0.19053856,
similarity=20.444647]
12: 3672106 Chris_Walker_(squash_player)
[pem_pos=4, final_scores=0.60158706,
final_local_scores=0.74781686, pem=-1.4941665,
similarity=34.53157]
13: 9316 England
[pem_pos=0, final_scores=0.55325145,
final_local_scores=0.5262059, pem=-0.19053856,
similarity=20.828924]
14: 5843419 France
[pem_pos=0, final_scores=0.73127794,
final_local_scores=0.76019335, pem=-0.16240834,
similarity=27.6586]
15: 1529068 Jonathon_Power
[pem_pos=0, final_scores=0.4976815,
final_local_scores=0.5808524, pem=0.0, similarity=21.423508]
16: 5042916 Canada
[pem_pos=0, final_scores=0.7799538,
final_local_scores=0.7949765, pem=-0.11662793,
similarity=28.447638]
17: 3672222 Ahmed_Barada
[pem_pos=0, final_scores=0.50354165,
final_local_scores=0.6261918, pem=0.0, similarity=22.776525]
18: 8087628 Egypt
[pem_pos=0, final_scores=0.5833258,
final_local_scores=0.60362, pem=-0.27480945,
similarity=23.597395]
19: 3451400 Amr_Shabana
[pem_pos=0, final_scores=0.6062175,
final_local_scores=0.73619026, pem=0.0, similarity=26.0591]
20: 8087628 Egypt
[pem_pos=0, final_scores=0.5654655,
final_local_scores=0.5841103, pem=-0.27480945,
similarity=23.015186]
21: 3671720 John_White_(squash_player)
[pem_pos=4, final_scores=0.40071768,
final_local_scores=0.50914365, pem=-2.7962391,
similarity=34.489956]
22: 4689264 Australia
[pem_pos=0, final_scores=0.7185417,
final_local_scores=0.72547764, pem=-0.17168477,
similarity=26.673061]
23: 7434625 Paul_Johnson_(squash_player)
[pem_pos=9, final_scores=0.52090436,
final_local_scores=0.65995765, pem=-2.962828,
similarity=39.896484]
24: 9316 England
[pem_pos=0, final_scores=0.56411844,
final_local_scores=0.5380764, pem=-0.19053856,
similarity=21.183165]
```

```
25: 9316 England
    [pem_pos=0, final_scores=0.58825356,
final_local_scores=0.56444037, pem=-0.19053856,
similarity=21.969917]
26: 23235 Pakistan
    [pem_pos=0, final_scores=0.6137348,
final_local_scores=0.6426545, pem=-0.21944217,
similarity=24.461164]
27: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5183652,
final_local_scores=0.5106174, pem=-0.22402541,
similarity=20.54584]
```

Sample: b'659_BASKETBALL'

BASKETBALL - [1 1 FORMULA SHELL 1] WIN GAME ONE IN
PHILIPPINES . MANILA 1996-08-28 Result of game one of the [2 2
Philippine Basketball Association 2] second conference finals
on Tuesday : [3 3 Formula Shell 3] beat [4 4 Alaska Milk 4]
85-82 (36-46) ([5 5 Formula Shell 5] leads best-of-seven
series 1-0)

```
1: 4092880 Shell_Turbo_Chargers
2: 364158 Philippine_Basketball_Association
3: 4092880 Shell_Turbo_Chargers
4: 4047853 Alaska_Aces_(PBA)
5: 4092880 Shell_Turbo_Chargers
1: 4092880 Shell_Turbo_Chargers
    [pem_pos=0, final_scores=0.44464672,
final_local_scores=0.37419975, pem=0.0, similarity=15.256579]
2: 364158 Philippine_Basketball_Association
    [pem_pos=0, final_scores=0.38256484,
final_local_scores=0.35919878, pem=0.0, similarity=14.808921]
3: 4092880 Shell_Turbo_Chargers
    [pem_pos=0, final_scores=0.65417296,
final_local_scores=0.60307527, pem=0.0, similarity=22.086681]
4: 4047853 Alaska_Aces_(PBA)
    [pem_pos=0, final_scores=0.5990769,
final_local_scores=0.58211696, pem=0.0, similarity=21.461245]
5: 4092880 Shell_Turbo_Chargers
    [pem_pos=0, final_scores=0.56386715,
final_local_scores=0.50443, pem=0.0, similarity=19.14291]
```

Sample: b'660_SOCCER'

SOCCER - [1 1 ISRAELI 1] FIRST DIVISION RESULTS . [2 2 JERUS
ALEM 2] 1996-08-28 Results of first division soccer matches
played over the weekend and Tuesday : Hapoel Kfar Sava 0 [3 3
Hapoel Zafirim Holon 3] 1 [4 4 Hapoel Tel Aviv 4] 1 [5 5 Mac
cabi Haifa 5] 3 [6 6 Hapoel Jerusalem 6] 0 [7 7 Hapoel Petah T
ikva 7] 3 [8 8 Hapoel Ironi Rishon 8] Lezion 3 [9 9 Hapoel Ta
ibe 9] 1 Hapoel Beit She'an 0 Hapoel Beit She'an 1 [10 10 Mac
cabi Petah Tikva 10] 0 [11 11 Betar Jerusalem 11] 3 [12 12 Ha
poel Haifa 12] 3 Maccabi [13 13 Tel Aviv 13] 1 Hapoel
Beersheva 2 [14 14 Maccabi Herzliya 14] 0

```
1: 9282173 Israel
2: 16043 Jerusalem
3: 10400096 Hapoel_Tzafririm_Holon_F.C.
4: 2301245 Hapoel_Tel_Aviv_F.C.
5: 319872 Maccabi_Haifa_F.C.
6: 5834227 Hapoel_Jerusalem_F.C.
7: 5166496 Hapoel_Petah_Tikva_F.C.
8: 5897396 Hapoel_Rishon_LeZion_F.C.
9: 19486919 Hapoel_Tayibe_F.C.
10: 893539 Maccabi_Petah_Tikva_F.C.
11: 1755537 Beitar_Jerusalem_F.C.
12: 2869548 Hapoel_Haifa_F.C.
13: 2301245 Hapoel_Tel_Aviv_F.C.
14: 5333420 Maccabi_Herzliya_F.C.

1: 9282173 Israel
    [pem_pos=0, final_scores=0.5712434,
final_local_scores=0.6319831, pem=-0.34790626,
similarity=24.841316]
2: 16043 Jerusalem
    [pem_pos=0, final_scores=0.4237569,
final_local_scores=0.49970657, pem=-0.13239029,
similarity=19.721912]
3: 10400096 Hapoel_Tzafririm_Holon_F.C.
    [pem_pos=0, final_scores=0.29868197,
final_local_scores=0.29947823, pem=0.0, similarity=13.026741]
4: 2301245 Hapoel_Tel_Aviv_F.C.
    [pem_pos=0, final_scores=0.26137692,
final_local_scores=0.19105047, pem=0.0, similarity=9.79104]
5: 319872 Maccabi_Haifa_F.C.
    [pem_pos=0, final_scores=0.31328315,
final_local_scores=0.2527114, pem=0.0, similarity=11.6311245]
6: 5834227 Hapoel_Jerusalem_F.C.
    [pem_pos=1, final_scores=0.26331437,
final_local_scores=0.2240023, pem=-0.3106352,
similarity=12.463672]
7: 5166496 Hapoel_Petah_Tikva_F.C.
    [pem_pos=0, final_scores=0.34213942,
final_local_scores=0.28992504, pem=0.0, similarity=12.741654]
8: 5897396 Hapoel_Rishon_LeZion_F.C.
    [pem_pos=0, final_scores=0.35464734,
final_local_scores=0.3195918, pem=0.0, similarity=13.626969]
9: 19486919 Hapoel_Tayibe_F.C.
    [pem_pos=0, final_scores=0.43268746,
final_local_scores=0.48352402, pem=0.0, similarity=18.519033]
10: 893539 Maccabi_Petah_Tikva_F.C.
    [pem_pos=0, final_scores=0.3543532,
final_local_scores=0.30056554, pem=0.0, similarity=13.059188]
11: 1755537 Beitar_Jerusalem_F.C.
    [pem_pos=0, final_scores=0.44673723,
final_local_scores=0.40103883, pem=0.0, similarity=16.057512]
12: 2869548 Hapoel_Haifa_F.C.
    [pem_pos=0, final_scores=0.35626572,
```

```
final_local_scores=0.3016624, pem=0.0, similarity=13.09192]
13: 2301245 Hapoel_Tel_Aviv_F.C.
    [pem_pos=6, final_scores=0.16337956,
final_local_scores=0.08400327, pem=-2.3020413,
similarity=19.115396]
14: 5333420 Maccabi_Herzliya_F.C.
    [pem_pos=0, final_scores=0.35143048,
final_local_scores=0.31937605, pem=0.0, similarity=13.620531]
```

Sample: b'661_TENNIS'

TENNIS - SELES HAS WALKOVER TO [1 1 U.S. OPEN 1] THIRD ROUND . [2 2 NEW YORK 2] 1996-08-28 Second seed and co-world number one [3 3 Monica Seles 3] advanced to the third round of the [4 4 U.S. Open Tennis Championships 4] without hitting a ball on Wednesday . [5 5 Seles 5] , the 1991 and 1992 champion who dropped just one game in her opening match , was scheduled to play [6 6 Laurence Courtois 6] of [7 7 Belgium 7] Wednesday night . But tournament officials announced about four-and-a-half hours before the match that [8 8 Courtois 8] had pulled out due to a left knee bone inflammation , moving [9 9 Seles 9] into the next round on a walkover .

```
1: 82729 US_Open_(tennis)
2: 645042 New_York_City
3: 300468 Monica_Seles
4: 82729 US_Open_(tennis)
5: 300468 Monica_Seles
6: 3004436 Laurence_Courtois
7: 3343 Belgium
8: 3004436 Laurence_Courtois
9: 300468 Monica_Seles
1: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.562908,
final_local_scores=0.564626, pem=-1.2370842,
similarity=27.666735]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.33470654,
final_local_scores=0.45431137, pem=-1.7906687,
similarity=27.385204]
3: 300468 Monica_Seles
    [pem_pos=0, final_scores=0.7965402,
final_local_scores=0.77519506, pem=0.0, similarity=27.22308]
4: 82729 US_Open_(tennis)
    [pem_pos=0, final_scores=0.6090272,
final_local_scores=0.61500424, pem=0.0, similarity=22.442667]
5: 300468 Monica_Seles
    [pem_pos=0, final_scores=0.6432561,
final_local_scores=0.6160956, pem=0.0, similarity=22.475235]
6: 3004436 Laurence_Courtois
    [pem_pos=0, final_scores=1.0749924,
final_local_scores=1.1463752, pem=0.0, similarity=38.29983]
7: 3343 Belgium
    [pem_pos=0, final_scores=0.6916609,
```

```
final_local_scores=0.798337, pem=-0.24180195,
similarity=29.228638]
8: 3004436 Laurence_Courtois
    [pem_pos=0, final_scores=0.7456015,
final_local_scores=0.80837387, pem=0.0, similarity=28.213202]
9: 300468 Monica_Seles
    [pem_pos=0, final_scores=0.62564117,
final_local_scores=0.5968539, pem=0.0, similarity=21.901024]
```

Sample: b'662_TENNIS'

TENNIS - SERIOUS [1 1 MEDVEDEV 1] IS HAVING FUN AGAIN .
Richard Finn [2 2 NEW YORK 2] 1996-08-28 Outspoken [3 3 Andrei Medvedev 3] exchanged his reputation as the clown prince of tennis on Wednesday for a new no-nonsense attitude that has made life on the courts fun again . " I think I 'm much more focused on what I have to do , and that 's playing tennis , " [4 4 Medvedev 4] said after routing [5 5 Frenchman 5] [6 6 Jean-Philippe Fleurian 6] 6-2 6-0 6-1 in the opening round of the [7 7 U.S. Open 7] . It was [8 8 Medvedev 8] 's sixth victory in a row after winning his first tournament of the year last week at the [9 9 Hamlet Cup 9] . " I realised this year , that without putting 99.9 percent of your mind into tennis , I do n't think you can successful , " said the 22-year-old [10 10 Medvedev 10] . " The whole day I 'm thinking about tennis . I felt that all the other things I was doing the years before , they were distracting me , they were not helping me at all . " For [11 11 Medvedev 11] that meant confining his post-match comments to tennis and not going off on tirades about peripheral issues such as the poor quality of food in the players lounge , an entertaining rant that took his mind off the task at hand . " I know what I 'm here for , " said [12 12 Medvedev 12] , who lost in the second round of the Open the last two years after reaching the quarters in 1993 , the same year he tried his hand as a restaurant critic . " I 'm not here to fight the press or talk about the food or entertain the people off the court . I 'm here to play tennis and to win . I have much less fun off the court . I have much more fun on the court , " he said . Just three years ago [13 13 Medvedev 13] was one of the world 's best , with a ranking of six after reaching the [14 14 French Open 14] semifinal and winning three tournaments . But [15 15 Medvedev 15] 's ranking slowly began to drop last year as he struggled with a wrist injury . The [16 16 Ukrainian 16] finally hit a low of 44th two months ago . " It 's somewhere where I would n't like to stay very long , " [17 17 Medvedev 17] said of his current ranking of 36 . " It 's a part of the penalty that I have to accept . " As part of his new businesslike approach , [18 18 Medvedev 18] hired [19 19 Australian 19] coach [20 20 Bob Brett 20] at the start of this year and the partnership is beginning to pay off . " At the beginning of the year we started from zero , " said [21 21 Medvedev 21] . " Winning in [22 22 Long Island 22] (last

week) was like winning for the first time . " While [23 23 Medvedev 23] 's 77-minute romp past [24 24 Fleurian 24] was rather ordinary , the fact that the two were playing each other was rather remarkable . In the original draw , [25 25 Medvedev 25] and [26 26 Fleurian 26] were slotted to play each other . When controversy forced the draw to be done over -- against odds of 151-to-1 -- [27 27 Medvedev 27] and [28 28 Fleurian 28] drew each other a second time . " When I saw the new draw I didn't have to change my preparation , " [29 29 Medvedev 29] said . " I think it 's destined that it turned out for me . "

```
1: 702849 Andriy_Medvedev
2: 645042 New_York_City
3: 702849 Andriy_Medvedev
4: 702849 Andriy_Medvedev
5: 5843419 France
6: 4026812 Jean-Philippe_Fleurian
7: 82729 US_Open_(tennis)
8: 702849 Andriy_Medvedev
9: 2809984 New_Haven_Open_at_Yale
10: 702849 Andriy_Medvedev
11: 702849 Andriy_Medvedev
12: 702849 Andriy_Medvedev
13: 702849 Andriy_Medvedev
14: 147724 French_Open
15: 702849 Andriy_Medvedev
16: 31750 Ukraine
17: 702849 Andriy_Medvedev
18: 702849 Andriy_Medvedev
19: 4689264 Australia
20: 26327557 Bob_Brett
21: 702849 Andriy_Medvedev
22: 2809984 New_Haven_Open_at_Yale
23: 702849 Andriy_Medvedev
24: 4026812 Jean-Philippe_Fleurian
25: 702849 Andriy_Medvedev
26: 4026812 Jean-Philippe_Fleurian
27: 702849 Andriy_Medvedev
28: 4026812 Jean-Philippe_Fleurian
29: 702849 Andriy_Medvedev
1: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=0.9075387,
final_local_scores=0.87095183, pem=0.0, similarity=30.080654]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.26134044,
final_local_scores=0.402389, pem=-1.7906687,
similarity=25.835737]
3: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=0.95343715,
final_local_scores=0.9206825, pem=0.0, similarity=31.564716]
4: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=1.0613644,
```

```
final_local_scores=1.0389831, pem=0.0, similarity=35.095043]
5: 5843419 France
    [pem_pos=1, final_scores=0.4627158,
final_local_scores=0.58410126, pem=-0.7184992,
similarity=25.42777]
6: 4026812 Jean-Philippe_Fleurian
    [pem_pos=0, final_scores=0.6729542,
final_local_scores=0.76693094, pem=0.0, similarity=26.976461]
7: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.47147304,
final_local_scores=0.50271773, pem=-1.2370842,
similarity=25.81927]
8: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=0.99466044,
final_local_scores=0.9661191, pem=0.0, similarity=32.920635]
9: 2809984 New_Haven_Open_at_Yale
    [pem_pos=0, final_scores=0.44197458,
final_local_scores=0.5027916, pem=0.0, similarity=19.094017]
10: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=0.9120911,
final_local_scores=0.87592465, pem=0.0, similarity=30.229052]
11: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=1.0584348,
final_local_scores=1.035783, pem=0.0, similarity=34.999542]
12: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=1.1073706,
final_local_scores=1.0892379, pem=0.0, similarity=36.594746]
13: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=1.0604129,
final_local_scores=1.0379438, pem=0.0, similarity=35.06403]
14: 147724 French_Open
    [pem_pos=0, final_scores=0.44990975,
final_local_scores=0.45331395, pem=-0.23693618,
similarity=18.906]
15: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=0.99744946,
final_local_scores=0.96916574, pem=0.0, similarity=33.01155]
16: 31750 Ukraine
    [pem_pos=1, final_scores=0.5016705,
final_local_scores=0.6531848, pem=-1.0458583,
similarity=29.26959]
17: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=1.0734515,
final_local_scores=1.0521865, pem=0.0, similarity=35.48906]
18: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=0.98881966,
final_local_scores=0.959739, pem=0.0, similarity=32.73024]
19: 4689264 Australia
    [pem_pos=0, final_scores=0.47315472,
final_local_scores=0.6347572, pem=-0.491009,
similarity=25.702318]
20: 26327557 Bob_Brett
```

```

[pem_pos=0, final_scores=0.38210464,
final_local_scores=0.5052856, pem=0.0, similarity=19.168444]
21: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=0.9992483,
final_local_scores=0.97113067, pem=0.0, similarity=33.07019]
22: 2809984 New_Haven_Open_at_Yale
    [pem_pos=7, final_scores=0.33999473,
final_local_scores=0.39139408, pem=-3.7117617,
similarity=35.95483]
23: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=1.0504072,
final_local_scores=1.027014, pem=0.0, similarity=34.737858]
24: 4026812 Jean-Philippe_Fleurian
    [pem_pos=0, final_scores=0.6802312,
final_local_scores=0.77487993, pem=0.0, similarity=27.213676]
25: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=1.0775054,
final_local_scores=1.0566146, pem=0.0, similarity=35.6212]
26: 4026812 Jean-Philippe_Fleurian
    [pem_pos=0, final_scores=0.7767729,
final_local_scores=0.8803372, pem=0.0, similarity=30.360731]
27: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=1.1088009,
final_local_scores=1.0908003, pem=0.0, similarity=36.64137]
28: 4026812 Jean-Philippe_Fleurian
    [pem_pos=0, final_scores=0.64771354,
final_local_scores=0.7393594, pem=0.0, similarity=26.153671]
29: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=0.9581533,
final_local_scores=0.92624074, pem=0.0, similarity=31.730585]

```

Sample: b'663_TENNIS'

TENNIS - WEDNESDAY 'S RESULTS FROM THE [1 1 U.S. OPEN 1] . [2 2 NEW YORK 2] 1996-08-28 Results of Wednesday 's matches in the [3 3 U.S. Open Tennis Championships 3] at the National Tennis Centre (prefix number denotes seeding) : Women 's singles , second round 15 - [4 4 Gabriela Sabatini 4] ([5 5 Argentina 5]) beat [6 6 Ann Grossman 6] ([7 7 U.S. 7]) 6-2 6 - 3 [8 8 Irina Spirlea 8] ([9 9 Romania 9]) beat Maria Jose Gaidano ([10 10 Argentina 10]) 6-1 6-2 8 - [11 11 Lindsay Davenport 11] ([12 12 U.S. 12]) beat Henrietta Nagyova ([13 13 Slovakia 13]) 6- 0 6-4 [14 14 Anne-Gaelle Sidot 14] ([15 15 France 15]) beat Wang Shi-Ting ([16 16 Taiwan 16]) 6-4 3-6 6-3 [17 17 Sandrine Testud 17] ([18 18 France 18]) beat Cristina Torrens-Valero ([19 19 Spain 19]) 6 - 2 6-1 Men 's singles , first round [20 20 Andrei Medvedev 20] ([21 21 Ukraine 21]) beat [22 22 Jean-Philippe Fleurian 22] ([23 23 France 23]) 6-2 6-0 6-1 David Nainkin ([24 24 South Africa 24]) beat 9 - [25 25 Wayne Ferreira 25] ([26 26 South Africa 26]) 6-4 6-4 2-6 7-5 [27 27 David Rikl 27] ([28 28 Czech Republic 28]) beat [29 29 Hicham Arazi 29] ([30 30 Morocco 30]) 6-4 7-5 6-2 [31 31 Andrea Gaudenzi 31] ([32 32 Italy 32])

ly 32]) beat [33 33 Shuzo Matsuoka 33] ([34 34 Japan 34])
7-6 (7-4) 6 - 2 6-3 Men 's singles , first round 17 - [35 35
Felix Mantilla 35] ([36 36 Spain 36]) beat [37 37 Fernando M
eligeni 37] ([38 38 Brazil 38]) 6-1 6 - 7 (2-7) 7-6
(7-5) 6-3 [39 39 Jonas Bjorkman 39] ([40 40 Sweden 40])
beat [41 41 Karol Kucera 41] ([42 42 Slovakia 42]) 6-2 5-7
7- 6 (7-3) 7-5 [43 43 Jan Kroslak 43] ([44 44 Slovakia 44]
) beat [45 45 Chris Woodruff 45] ([46 46 U.S. 46]) 2-6 6-4
3-6 6 - 2 7-6 (7-1) Women 's singles , second round [47 47 A
manda Coetzer 47] ([48 48 South Africa 48]) beat [49 49 Mar
iaan de Swardt 49] ([50 50 South Africa 50]) 6-2 7-5 [51 51
Linda Wild 51] ([52 52 U.S. 52]) beat [53 53 Kristie Boogert
53] ([54 54 Netherlands 54]) 5-7 6-3 6-3 [55 55 Kimberly Po
55] ([56 56 U.S. 56]) beat [57 57 Kristina Brandi 57] ([58
58 U.S. 58]) 6-1 6-4 [59 59 Helena Sukova 59] ([60 60 Czech
Republic 60]) beat [61 61 Paola Suarez 61] ([62 62 Argentina
62]) 6- 4 7-6 (7-2) Women 's singles , second round 2 - [63
63 Monica Seles 63] ([64 64 U.S. 64]) beat [65 65 Laurence C
ourtois 65] ([66 66 Belgium 66]) by walkover (knee
injury) [67 67 Dally Randriantefy 67] ([68 68 Madagascar 68]
) beat Jane Chi ([69 69 U.S. 69]) 6-3 6-1 [70 70 Ines Gorro
chategui 70] ([71 71 Argentina 71]) beat [72 72 Aleksandra O
lsza 72] ([73 73 Poland 73]) 6 - 1 6-1 Men 's singles ,
first round 12 - [74 74 Todd Martin 74] ([75 75 U.S. 75])
beat [76 76 Younes El Aynaoui 76] ([77 77 Morocco 77]) 6-3
6-2 4-6 6-4 [78 78 Sjeng Schalken 78] ([79 79 Netherlands 79]
) beat [80 80 Gilbert Schaller 80] ([81 81 Austria 81]) 6-
3 6-4 6-7 (6-8) 6-3 Men 's singles , first round [82 82 Mic
hael Tebbutt 82] ([83 83 Australia 83]) beat [84 84 Richey R
eneberg 84] ([85 85 U.S. 85]) 3-6 6-1 3-6 7-5 6-3 [86 86 Pa
ul Haarhuis 86] ([87 87 Netherlands 87]) beat [88 88 Michael
Joyce 88] ([89 89 U.S 89]) 6-7 (5-7) 7-6 (8-6) 1-6 6-2
6-2 Women 's singles , second round [90 90 Barbara Rittner 90]
([91 91 Germany 91]) beat 13 - [92 92 Brenda Schultz-
McCarthy 92] ([93 93 Netherlands 93]) 6-2 6-1 Men 's
singles , first round [94 94 Guy Forget 94] ([95 95 France 95]
) beat [96 96 Grant Stafford 96] ([97 97 South Africa
97]) 3-6 2-6 6-4 7-6 (7-2) 6-3 (End first round) Women
's singles , second round [98 98 Lisa Raymond 98] ([99 99 U.
S. 99]) beat [100 100 Sarah Pitkowski 100] ([101 101 France
101]) 6-2 6-0 [102 102 Asa Carlsson 102] ([103 103 Sweden 10
3]) beat [104 104 Barbara Schett 104] ([105 105 Austria 105]
) 6-2 3-1 retired (Thigh injury)

- 1: 82729 US_Open_(tennis)
- 2: 645042 New_York_City
- 3: 82729 US_Open_(tennis)
- 4: 782187 Gabriela_Sabatini
- 5: 18951905 Argentina
- 6: 15097221 Ann_Grossman
- 7: 3434750 United_States
- 8: 4038084 Irina_Spirlea
- 9: 25445 Romania

10: 18951905 Argentina
11: 52828 Lindsay_Davenport
12: 3434750 United_States
13: 26830 Slovakia
14: 4450757 Anne-Gaëlle_Sidot
15: 5843419 France
16: 25734 Taiwan
17: 3127711 Sandrine_Testud
18: 5843419 France
19: 26667 Spain
20: 702849 Andriy_Medvedev
21: 31750 Ukraine
22: 4026812 Jean-Philippe_Fleurian
23: 5843419 France
24: 17416221 South_Africa
25: 1723516 Wayne_Ferreira
26: 17416221 South_Africa
27: 10250905 David_Rikl
28: 5321 Czech_Republic
29: 2508144 Hicham_Arazi
30: 19291 Morocco
31: 3929637 Andrea_Gaudenzi
32: 14532 Italy
33: 2575718 Shuzo_Matsuoka
34: 15573 Japan
35: 1723559 Félix_Mantilla_Botella
36: 26667 Spain
37: 2989127 Fernando_Meligeni
38: 3383 Brazil
39: 1964402 Jonas_Björkman
40: 5058739 Sweden
41: 3039956 Karol_Kučera
42: 26830 Slovakia
43: 3930155 Ján_Krošlák
44: 26830 Slovakia
45: 2515285 Chris_Woodruff
46: 3434750 United_States
47: 1951841 Amanda_Coetzer
48: 17416221 South_Africa
49: 10119577 Mariaan_de_Swardt
50: 17416221 South_Africa
51: 15442936 Linda_Wild
52: 3434750 United_States
53: 2527532 Kristie_Boogert
54: 21148 Netherlands
55: 6426517 Kimberly_Po
56: 3434750 United_States
57: 1301702 Kristina_Brandi
58: 3434750 United_States
59: 1920271 Helena_Suková
60: 5321 Czech_Republic
61: 705473 Paola_Suárez

62: 18951905 Argentina
63: 300468 Monica_Seles
64: 3434750 United_States
65: 3004436 Laurence_Courtois
66: 3343 Belgium
67: 2620429 Dally_Randriantefy
68: 18964 Madagascar
69: 3434750 United_States
70: 6501697 Inés_Gorrochategui
71: 18951905 Argentina
72: 22207361 Aleksandra_Olsza
73: 22936 Poland
74: 871725 Todd_Martin
75: 3434750 United_States
76: 2090929 Younes_El_Aynaoui
77: 19291 Morocco
78: 2493300 Sjeng_Schalken
79: 21148 Netherlands
80: 20856583 Gilbert_Schaller
81: 26964606 Austria
82: 13355036 Michael_Tebbutt
83: 4689264 Australia
84: 871728 Richey_Reneberg
85: 3434750 United_States
86: 2186352 Paul_Haarhuis
87: 21148 Netherlands
88: 4026170 Michael_Joyce_(tennis)
89: 3434750 United_States
90: 12668474 Barbara_Rittner
91: 11867 Germany
92: 3484709 Brenda_Schultz-McCarthy
93: 21148 Netherlands
94: 1723582 Guy_Forget
95: 5843419 France
96: 4026564 Grant_Stafford
97: 17416221 South_Africa
98: 2199290 Lisa_Raymond
99: 3434750 United_States
100: 3127732 Sarah_Pitkowski-Malcor
101: 5843419 France
102: 4038174 Åsa_Svensson
103: 5058739 Sweden
104: 2516435 Barbara_Schett
105: 26964606 Austria
1: 82729 US_Open_(tennis)
 [pem_pos=1, final_scores=0.52107763,
final_local_scores=0.5354991, pem=-1.2370842,
similarity=26.79753]
2: 645042 New_York_City
 [pem_pos=1, final_scores=0.42656428,
final_local_scores=0.46009672, pem=-1.7906687,
similarity=27.557852]

```
3: 82729 US_Open_(tennis)
    [pem_pos=0, final_scores=0.5590539,
final_local_scores=0.5769824, pem=0.0, similarity=21.308018]
4: 782187 Gabriela_Sabatini
    [pem_pos=0, final_scores=0.8307782,
final_local_scores=0.91627663, pem=0.0, similarity=31.433237]
5: 18951905 Argentina
    [pem_pos=0, final_scores=0.703092,
final_local_scores=0.72351074, pem=-0.30003807,
similarity=27.31237]
6: 15097221 Ann_Grossman
    [pem_pos=0, final_scores=0.7921334,
final_local_scores=0.92958647, pem=0.0, similarity=31.83043]
7: 3434750 United_States
    [pem_pos=0, final_scores=0.52709407,
final_local_scores=0.5320869, pem=-0.14075686,
similarity=20.733706]
8: 4038084 Irina_Spirlea
    [pem_pos=0, final_scores=0.9002605,
final_local_scores=1.0210471, pem=0.0, similarity=34.559795]
9: 25445 Romania
    [pem_pos=0, final_scores=0.76626927,
final_local_scores=0.79582673, pem=-0.21241479,
similarity=28.993914]
10: 18951905 Argentina
    [pem_pos=0, final_scores=0.67252535,
final_local_scores=0.69012135, pem=-0.30003807,
similarity=26.315964]
11: 52828 Lindsay_Davenport
    [pem_pos=0, final_scores=0.8390799,
final_local_scores=0.9107904, pem=0.0, similarity=31.269516]
12: 3434750 United_States
    [pem_pos=0, final_scores=0.49988675,
final_local_scores=0.5023671, pem=-0.14075686,
similarity=19.846806]
13: 26830 Slovakia
    [pem_pos=0, final_scores=0.6431963,
final_local_scores=0.6517271, pem=-0.36602598,
similarity=25.529057]
14: 4450757 Anne-Gaëlle_Sidot
    [pem_pos=0, final_scores=0.5451494,
final_local_scores=0.66505504, pem=0.0, similarity=23.936281]
15: 5843419 France
    [pem_pos=0, final_scores=0.74788225,
final_local_scores=0.7600256, pem=-0.16240834,
similarity=27.653595]
16: 25734 Taiwan
    [pem_pos=0, final_scores=0.38251275,
final_local_scores=0.4231066, pem=-0.077886656,
similarity=17.139618]
17: 3127711 Sandrine_Testud
    [pem_pos=0, final_scores=0.7027311,
```

```
final_local_scores=0.7990674, pem=0.0, similarity=27.935478]
18: 5843419 France
    [pem_pos=0, final_scores=0.73495626,
final_local_scores=0.74590594, pem=-0.16240834,
similarity=27.232235]
19: 26667 Spain
    [pem_pos=0, final_scores=0.67327636,
final_local_scores=0.69144744, pem=-0.18857051,
similarity=25.74936]
20: 702849 Andriy_Medvedev
    [pem_pos=0, final_scores=0.87425715,
final_local_scores=0.9642646, pem=0.0, similarity=32.86529]
21: 31750 Ukraine
    [pem_pos=0, final_scores=0.6113886,
final_local_scores=0.641, pem=-0.35407752,
similarity=25.143959]
22: 4026812 Jean-Philippe_Fleurian
    [pem_pos=0, final_scores=0.63000816,
final_local_scores=0.7298066, pem=0.0, similarity=25.868597]
23: 5843419 France
    [pem_pos=0, final_scores=0.7331364,
final_local_scores=0.743918, pem=-0.16240834,
similarity=27.17291]
24: 17416221 South_Africa
    [pem_pos=0, final_scores=0.6306061,
final_local_scores=0.679432, pem=-0.26251382,
similarity=25.79291]
25: 1723516 Wayne_Ferreira
    [pem_pos=0, final_scores=0.70135003,
final_local_scores=0.8000446, pem=0.0, similarity=27.96464]
26: 17416221 South_Africa
    [pem_pos=0, final_scores=0.52978265,
final_local_scores=0.5692977, pem=-0.26251382,
similarity=22.506283]
27: 10250905 David_Rikl
    [pem_pos=0, final_scores=0.6548279,
final_local_scores=0.76369643, pem=0.0, similarity=26.879938]
28: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.57239264,
final_local_scores=0.57687926, pem=-0.27794555,
similarity=22.816452]
29: 2508144 Hicham_Arazi
    [pem_pos=0, final_scores=0.6253547,
final_local_scores=0.72171587, pem=0.0, similarity=25.627153]
30: 19291 Morocco
    [pem_pos=0, final_scores=0.61557144,
final_local_scores=0.6532225, pem=-0.29598162,
similarity=25.192768]
31: 3929637 Andrea_Gaudenzi
    [pem_pos=0, final_scores=0.83468,
final_local_scores=0.959881, pem=0.0, similarity=32.734478]
32: 14532 Italy
```

```
[pem_pos=0, final_scores=0.7387628,
final_local_scores=0.7867791, pem=-0.20612453,
similarity=28.689709]
33: 2575718 Shuzo_Matsuoka
    [pem_pos=0, final_scores=0.68778366,
final_local_scores=0.811718, pem=0.0, similarity=28.312996]
34: 15573 Japan
    [pem_pos=0, final_scores=0.71801066,
final_local_scores=0.7710925, pem=-0.13549899,
similarity=27.837515]
35: 1723559 Félix_Mantilla_Botella
    [pem_pos=0, final_scores=0.6936485,
final_local_scores=0.7695474, pem=0.0, similarity=27.054543]
36: 26667 Spain
    [pem_pos=0, final_scores=0.65722674,
final_local_scores=0.6739157, pem=-0.18857051,
similarity=25.226177]
37: 2989127 Fernando_Meligeni
    [pem_pos=0, final_scores=0.5463482,
final_local_scores=0.5944453, pem=0.0, similarity=21.829145]
38: 3383 Brazil
    [pem_pos=0, final_scores=0.8267496,
final_local_scores=0.86810297, pem=-0.21313342,
similarity=31.154692]
39: 1964402 Jonas_Björkman
    [pem_pos=0, final_scores=0.70418113,
final_local_scores=0.8051577, pem=0.0, similarity=28.117226]
40: 5058739 Sweden
    [pem_pos=0, final_scores=0.7956056,
final_local_scores=0.82969916, pem=-0.2135508,
similarity=30.010912]
41: 3039956 Karol_Kučera
    [pem_pos=0, final_scores=0.7680635,
final_local_scores=0.85598874, pem=0.0, similarity=29.634125]
42: 26830 Slovakia
    [pem_pos=0, final_scores=0.6702718,
final_local_scores=0.681303, pem=-0.36602598,
similarity=26.411661]
43: 3930155 Ján_Krošlák
    [pem_pos=0, final_scores=0.5568289,
final_local_scores=0.6267927, pem=0.0, similarity=22.794458]
44: 26830 Slovakia
    [pem_pos=0, final_scores=0.6735423,
final_local_scores=0.68487555, pem=-0.36602598,
similarity=26.518272]
45: 2515285 Chris_Woodruff
    [pem_pos=0, final_scores=0.6779351,
final_local_scores=0.78696936, pem=0.0, similarity=27.574448]
46: 3434750 United_States
    [pem_pos=0, final_scores=0.5248396,
final_local_scores=0.5296243, pem=-0.14075686,
similarity=20.660217]
```

```
47: 1951841 Amanda_Coetzer
    [pem_pos=0, final_scores=0.7808116,
final_local_scores=0.86564523, pem=0.0, similarity=29.922295]
48: 17416221 South_Africa
    [pem_pos=0, final_scores=0.6346849,
final_local_scores=0.6838874, pem=-0.26251382,
similarity=25.925869]
49: 10119577 Mariaan_de_Swardt
    [pem_pos=0, final_scores=0.5808864,
final_local_scores=0.64927715, pem=0.0, similarity=23.465439]
50: 17416221 South_Africa
    [pem_pos=0, final_scores=0.5882027,
final_local_scores=0.63311267, pem=-0.26251382,
similarity=24.410648]
51: 15442936 Linda_Wild
    [pem_pos=0, final_scores=0.6643047,
final_local_scores=0.7828716, pem=0.0, similarity=27.452164]
52: 3434750 United_States
    [pem_pos=0, final_scores=0.48952907,
final_local_scores=0.49105293, pem=-0.14075686,
similarity=19.50917]
53: 2527532 Kristie_Boogert
    [pem_pos=0, final_scores=0.5153367,
final_local_scores=0.64092374, pem=0.0, similarity=23.216156]
54: 21148 Netherlands
    [pem_pos=0, final_scores=0.6423682,
final_local_scores=0.65052587, pem=-0.24940795,
similarity=24.859022]
55: 6426517 Kimberly_Po
    [pem_pos=0, final_scores=0.54959255,
final_local_scores=0.65194637, pem=0.0, similarity=23.545094]
56: 3434750 United_States
    [pem_pos=0, final_scores=0.46738833,
final_local_scores=0.4668675, pem=-0.14075686,
similarity=18.787428]
57: 1301702 Kristina_Brandi
    [pem_pos=0, final_scores=0.6163533,
final_local_scores=0.7064412, pem=0.0, similarity=25.171328]
58: 3434750 United_States
    [pem_pos=0, final_scores=0.49724752,
final_local_scores=0.49948412, pem=-0.14075686,
similarity=19.760775]
59: 1920271 Helena_Suková
    [pem_pos=0, final_scores=0.7984919,
final_local_scores=0.886354, pem=0.0, similarity=30.540285]
60: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.5966422,
final_local_scores=0.6033682, pem=-0.27794555,
similarity=23.606936]
61: 705473 Paola_Suárez
    [pem_pos=0, final_scores=0.74945825,
final_local_scores=0.82388395, pem=0.0, similarity=28.676054]
```

62: 18951905 Argentina
[pem_pos=0, final_scores=0.6818126,
final_local_scores=0.70026624, pem=-0.30003807,
similarity=26.618708]
63: 300468 Monica_Seles
[pem_pos=0, final_scores=0.89229983,
final_local_scores=0.9553333, pem=0.0, similarity=32.598763]
64: 3434750 United_States
[pem_pos=0, final_scores=0.49680012,
final_local_scores=0.49899536, pem=-0.14075686,
similarity=19.74619]
65: 3004436 Laurence_Courtois
[pem_pos=0, final_scores=0.9334745,
final_local_scores=1.0614548, pem=0.0, similarity=35.76564]
66: 3343 Belgium
[pem_pos=0, final_scores=0.6844164,
final_local_scores=0.68164617, pem=-0.24180195,
similarity=25.746351]
67: 2620429 Dally_Randriantefy
[pem_pos=0, final_scores=0.748564,
final_local_scores=0.90103906, pem=0.0, similarity=30.978516]
68: 18964 Madagascar
[pem_pos=0, final_scores=0.7006686,
final_local_scores=0.7956386, pem=-0.30672407,
similarity=29.501171]
69: 3434750 United_States
[pem_pos=0, final_scores=0.51709783,
final_local_scores=0.5211676, pem=-0.14075686,
similarity=20.40785]
70: 6501697 Inés_Gorrochategui
[pem_pos=0, final_scores=0.54464895,
final_local_scores=0.67633796, pem=0.0, similarity=24.272987]
71: 18951905 Argentina
[pem_pos=0, final_scores=0.64293665,
final_local_scores=0.6578002, pem=-0.30003807,
similarity=25.351437]
72: 22207361 Aleksandra_Olsza
[pem_pos=0, final_scores=0.6936299,
final_local_scores=0.85000783, pem=0.0, similarity=29.455643]
73: 22936 Poland
[pem_pos=0, final_scores=0.59848994,
final_local_scores=0.6086749, pem=-0.15333635,
similarity=23.087652]
74: 871725 Todd_Martin
[pem_pos=0, final_scores=0.82978886,
final_local_scores=0.907595, pem=0.0, similarity=31.174158]
75: 3434750 United_States
[pem_pos=0, final_scores=0.5115028,
final_local_scores=0.51505584, pem=-0.14075686,
similarity=20.225466]
76: 2090929 Younes_El_Aynaoui
[pem_pos=0, final_scores=0.67644614,

```
final_local_scores=0.7527408, pem=0.0, similarity=26.553]
77: 19291 Morocco
    [pem_pos=0, final_scores=0.6430611,
final_local_scores=0.6832507, pem=-0.29598162,
similarity=26.088871]
78: 2493300 Sjeng_Schalken
    [pem_pos=0, final_scores=0.80084914,
final_local_scores=0.8960703, pem=0.0, similarity=30.830236]
79: 21148 Netherlands
    [pem_pos=0, final_scores=0.6478297,
final_local_scores=0.6564917, pem=-0.24940795,
similarity=25.037054]
80: 20856583 Gilbert_Schaller
    [pem_pos=0, final_scores=0.86280245,
final_local_scores=0.9927661, pem=0.0, similarity=33.71583]
81: 26964606 Austria
    [pem_pos=0, final_scores=0.666938,
final_local_scores=0.67652595, pem=-0.25262418,
similarity=25.652407]
82: 13355036 Michael_Tebbutt
    [pem_pos=0, final_scores=0.5977246,
final_local_scores=0.71893513, pem=0.0, similarity=25.54417]
83: 4689264 Australia
    [pem_pos=0, final_scores=0.6928719,
final_local_scores=0.7300677, pem=-0.17168477,
similarity=26.810038]
84: 871728 Richey_Reneberg
    [pem_pos=0, final_scores=0.7321398,
final_local_scores=0.8379515, pem=0.0, similarity=29.095856]
85: 3434750 United_States
    [pem_pos=0, final_scores=0.51306576,
final_local_scores=0.51676315, pem=-0.14075686,
similarity=20.276415]
86: 2186352 Paul_Haarhuis
    [pem_pos=0, final_scores=0.5677226,
final_local_scores=0.66685754, pem=0.0, similarity=23.990072]
87: 21148 Netherlands
    [pem_pos=0, final_scores=0.6444522,
final_local_scores=0.6528023, pem=-0.24940795,
similarity=24.926954]
88: 4026170 Michael_Joyce_(tennis)
    [pem_pos=1, final_scores=0.4897303,
final_local_scores=0.5745407, pem=-0.41227743,
similarity=23.47718]
89: 3434750 United_States
    [pem_pos=0, final_scores=0.43817246,
final_local_scores=0.43495357, pem=0.0, similarity=17.069595]
90: 12668474 Barbara_Rittner
    [pem_pos=0, final_scores=0.59346765,
final_local_scores=0.68521583, pem=0.0, similarity=24.53792]
91: 11867 Germany
    [pem_pos=0, final_scores=0.82473737,
```

```
final_local_scores=0.8530275, pem=-0.15529114,  
similarity=30.390253]  
92: 3484709 Brenda_Schultz-McCarthy  
    [pem_pos=0, final_scores=0.7085568,  
final_local_scores=0.8154715, pem=0.0, similarity=28.425009]  
93: 21148 Netherlands  
    [pem_pos=0, final_scores=0.52253985,  
final_local_scores=0.51963156, pem=-0.24940795,  
similarity=20.952875]  
94: 1723582 Guy_Forget  
    [pem_pos=0, final_scores=0.72226757,  
final_local_scores=0.8033693, pem=0.0, similarity=28.063854]  
95: 5843419 France  
    [pem_pos=0, final_scores=0.78434926,  
final_local_scores=0.7998602, pem=-0.16240834,  
similarity=28.842337]  
96: 4026564 Grant_Stafford  
    [pem_pos=0, final_scores=0.9222253,  
final_local_scores=1.0493574, pem=0.0, similarity=35.40463]  
97: 17416221 South_Africa  
    [pem_pos=0, final_scores=0.60430765,  
final_local_scores=0.65070486, pem=-0.26251382,  
similarity=24.935635]  
98: 2199290 Lisa_Raymond  
    [pem_pos=0, final_scores=0.71623486,  
final_local_scores=0.8004847, pem=0.0, similarity=27.977774]  
99: 3434750 United_States  
    [pem_pos=0, final_scores=0.4863488,  
final_local_scores=0.48757893, pem=-0.14075686,  
similarity=19.405499]  
100: 3127732 Sarah_Pitkowski-Malcor  
    [pem_pos=0, final_scores=0.67746264,  
final_local_scores=0.7860673, pem=0.0, similarity=27.54753]  
101: 5843419 France  
    [pem_pos=0, final_scores=0.70147026,  
final_local_scores=0.7093276, pem=-0.16240834,  
similarity=26.140663]  
102: 4038174 Åsa_Svensson  
    [pem_pos=0, final_scores=0.8509303,  
final_local_scores=0.9896223, pem=0.0, similarity=33.622017]  
103: 5058739 Sweden  
    [pem_pos=0, final_scores=0.77741057,  
final_local_scores=0.8098239, pem=-0.2135508,  
similarity=29.417795]  
104: 2516435 Barbara_Schett  
    [pem_pos=0, final_scores=0.92398447,  
final_local_scores=1.026586, pem=0.0, similarity=34.72509]  
105: 26964606 Austria  
    [pem_pos=0, final_scores=0.68152016,  
final_local_scores=0.6924547, pem=-0.25262418,  
similarity=26.127752]
```

Sample: b'664_BASEBALL'

BASEBALL - [1 1 MAJOR LEAGUE 1] STANDINGS AFTER TUESDAY 'S GAMES . [2 2 NEW YORK 2] 1996-08-28 [3 3 Major League Baseball 3] standings after games played on Tuesday (tabulate under won , lost , winning percentage and games behind) : [4 4 AMERICAN LEAGUE EASTERN DIVISION 4] W L PCT GB [1 1 NEW YORK 1] 74 57 .565 - [5 5 BALTIMORE 5] 70 61 .534 4 [6 6 BOSTON 6] 68 65 .511 7 [2 2 TORONTO 2] 62 71 .466 13 [7 7 DETROIT 7] 47 85 .356 27 1/2 [8 8 CENTRAL DIVISION 8] [9 9 CLEVELAND 9] 79 53 .598 - 3 CHICAGO 70 64 .522 10 [10 10 MINNESOTA 10] 66 66 . 500 13 [11 11 MILWAUKEE 11] 64 69 .481 15 1/2 [12 12 KANSAS CITY 12] 60 73 .451 19 1/2 [13 13 WESTERN DIVISION 13] [14 14 TEXAS 14] 75 57 .568 - [15 15 SEATTLE 15] 68 63 .519 6 1/2 [16 16 OAKLAND 16] 63 72 .467 13 1/2 4 CALIFORNIA 61 71 .462 14 WEDNESDAY , AUGUST 28TH SCHEDULE [17 17 CLEVELAND 17] AT [18 18 DETROIT 18] [19 19 MILWAUKEE 19] AT 5 CHICAGO [20 20 OAKLAND 20] AT [21 21 BALTIMORE 21] [22 22 MINNESOTA 22] AT [23 23 TORONTO 23] [24 24 TEXAS 24] AT [25 25 KANSAS CITY 25] [26 26 BOSTON 26] AT 6 CALIFORNIA [27 27 NEW YORK 27] AT [28 28 SEATTLE 28] [29 29 NATIONAL LEAGUE EASTERN DIVISION 29] W L PCT GB [30 30 ATLANTA 30] 81 49 .623 - [31 31 MONTREAL 31] 70 60 .538 11 7 FLORIDA 62 70 .470 20 [32 32 NEW YORK 32] 59 73 .447 23 [33 33 PHILADELPHIA 33] 54 79 .406 28 1/2 [34 34 CENTRAL DIVISION 34] [35 35 HOUSTON 35] 71 62 .534 - ST LOUIS 69 63 .523 1 1/2 [36 36 CINCINNATI 36] 65 66 .496 5 8 CHICAGO 64 65 .496 5 [37 37 PITTSBURGH 37] 56 75 .427 14 [38 38 WESTERN DIVISION 38] [39 39 SAN DIEGO 39] 73 60 .549 - 9 LOS ANGELES 71 60 .542 1 [40 40 COLORADO 40] 69 64 .519 4 [41 41 SAN FRANCISCO 41] 56 74 . 431 15 1/2 WEDNESDAY , AUGUST 28TH SCHEDULE [42 42 CINCINNATI 42] AT [43 43 COLORADO 43] 10 LOS ANGELES AT [44 44 MONTREAL 44] [45 45 ATLANTA 45] AT [46 46 PITTSBURGH 46] [47 47 SAN DIEGO 47] AT [48 48 NEW YORK 48] 11 CHICAGO AT [49 49 HOUSTON 49] 12 FLORIDA AT ST LOUIS [50 50 PHILADELPHIA 50] AT [51 51 SAN FRANCISCO 51]

1: 38776 Major_League_Baseball
2: 645042 New_York_City
3: 38776 Major_League_Baseball
4: 2004582 American_League_East
5: 4340 Baltimore_Orioles
6: 4339 Boston_Red_Sox
7: 8579 Detroit_Tigers
8: 2456377 National_League_Central
9: 6652 Cleveland_Indians
10: 20050 Minnesota_Twins
11: 20152 Milwaukee_Brewers
12: 16883 Kansas_City_Royals
13: 2456474 National_League_West
14: 30857 Texas_Rangers_(baseball)
15: 27660 Seattle_Mariners
16: 22523 Oakland_Athletics
17: 6652 Cleveland_Indians
18: 8579 Detroit_Tigers

```
19: 20152 Milwaukee_Brewers
20: 22523 Oakland_Athletics
21: 4340 Baltimore_Orioles
22: 20050 Minnesota_Twins
23: 64646 Toronto
24: 30857 Texas_Rangers_(baseball)
25: 16883 Kansas_City_Royals
26: 4339 Boston_Red_Sox
27: 4848143 New_York_Yankees
28: 27660 Seattle_Mariners
29: 2456294 National_League_East
30: 2140 Atlanta_Braves
31: 20153 Montreal_Expos
32: 4848143 New_York_Yankees
33: 23741 Philadelphia_Phillies
34: 2456377 National_League_Central
35: 13894 Houston_Astros
36: 6671 Cincinnati_Reds
37: 24270 Pittsburgh_Pirates
38: 2456474 National_League_West
39: 28417 San_Diego_Padres
40: 6669 Colorado_Rockies
41: 28416 San_Francisco_Giants
42: 6671 Cincinnati_Reds
43: 6669 Colorado_Rockies
44: 7954681 Montreal
45: 2140 Atlanta_Braves
46: 24270 Pittsburgh_Pirates
47: 28417 San_Diego_Padres
48: 4848143 New_York_Yankees
49: 13894 Houston_Astros
50: 23741 Philadelphia_Phillies
51: 28416 San_Francisco_Giants
1: 645042 New_York_City
    [winner: 4848143 New_York_Yankees, pem_pos=11,
final_scores=0.242411, final_local_scores=0.15135652,
pem=-5.0692368, similarity=36.17379]
    gt: 645042 New_York_City, pem_pos=1,
final_scores=0.08078097, final_local_scores=0.11354381,
pem=-1.7906687, similarity=17.21602 ]
2: 64646 Toronto
    [winner: 30847 Toronto_Blue_Jays, pem_pos=6,
final_scores=0.27590877, final_local_scores=0.19582665,
pem=-4.422637, similarity=33.98456]
    gt: 64646 Toronto, pem_pos=0, final_scores=-0.0333803,
final_local_scores=-0.017672367, pem=-0.12291032,
similarity=4.2307396 ]
3: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.72610265, final_local_scores=-0.81709015,
pem=-4.591283, similarity=4.6742215]
    RECALL MISS: 5945 Chicago_White_Sox]
```

4: 1360083 Los_Angeles_Angels_ofAnaheim
[winner: 5407 California, pem_pos=0,
final_scores=-0.88845545, final_local_scores=-0.9302802,
pem=-0.08662026, similarity=-23.20065]
RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]

5: 5945 Chicago_White_Sox
[winner: 6886 Chicago, pem_pos=0,
final_scores=-0.64485335, final_local_scores=-0.70634204,
pem=-0.20658566, similarity=-15.865499]
RECALL MISS: 5945 Chicago_White_Sox]

6: 1360083 Los_Angeles_Angels_ofAnaheim
[winner: 5407 California, pem_pos=0,
final_scores=-0.7710466, final_local_scores=-0.8020289,
pem=-0.08662026, similarity=-19.373373]
RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]

7: 11278 Miami_Marlins
[winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.74743205, final_local_scores=-0.8189992,
pem=-4.620281, similarity=4.7749505]
RECALL MISS: 11278 Miami_Marlins]

8: 5945 Chicago_White_Sox
[winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.75830674, final_local_scores=-0.8522682,
pem=-4.591283, similarity=3.624439]
RECALL MISS: 5945 Chicago_White_Sox]

9: 1360083 Los_Angeles_Angels_ofAnaheim
[winner: 18213 Los_Angeles_Dodgers, pem_pos=15,
final_scores=-0.13228038, final_local_scores=-0.2595632,
pem=-4.7524934, similarity=22.188627]
RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]

10: 1360083 Los_Angeles_Angels_ofAnaheim
[winner: 18110 Los_Angeles, pem_pos=0,
final_scores=-0.06668638, final_local_scores=-0.05934968,
pem=-0.100652285, similarity=2.865963]
RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]

11: 5945 Chicago_White_Sox
[winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.63466513, final_local_scores=-0.71720856,
pem=-4.591283, similarity=7.6548886]
RECALL MISS: 5945 Chicago_White_Sox]

12: 11278 Miami_Marlins
[winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.6881856, final_local_scores=-0.7542814,
pem=-4.620281, similarity=6.706258]
RECALL MISS: 11278 Miami_Marlins]

1: 38776 Major_League_Baseball
[pem_pos=0, final_scores=0.58339113,
final_local_scores=0.53491277, pem=-0.47792417,
similarity=22.651602]

2: 645042 New_York_City
[pem_pos=1, final_scores=0.3634528,
final_local_scores=0.42221528, pem=-1.7906687,

```
similarity=26.427393]
3: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.54097766,
final_local_scores=0.4885825, pem=0.0, similarity=18.669989]
4: 2004582 American_League_East
    [pem_pos=0, final_scores=0.47077852,
final_local_scores=0.46921486, pem=0.0, similarity=18.09202]
5: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.2419118,
final_local_scores=1.2411897, pem=-4.046561,
similarity=63.135117]
6: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.1909184,
final_local_scores=1.1891294, pem=-4.418931,
similarity=63.60654]
7: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.2492768,
final_local_scores=1.2497087, pem=-4.1401463,
similarity=63.898273]
8: 2456377 National_League_Central
    [pem_pos=0, final_scores=0.5747995,
final_local_scores=0.59315115, pem=-0.98951703,
similarity=27.171673]
9: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.2986462,
final_local_scores=1.3025045, pem=-4.191953,
similarity=65.75554]
10: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=1.1490009,
final_local_scores=1.1422405, pem=-4.2906566,
similarity=61.509705]
11: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.53822607,
final_local_scores=0.47745818, pem=-3.40205,
similarity=36.838894]
12: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.71780777,
final_local_scores=0.6739278, pem=-3.5915844,
similarity=43.73266]
13: 2456474 National_League_West
    [pem_pos=1, final_scores=0.5641074,
final_local_scores=0.56917536, pem=-1.2952092,
similarity=28.11859]
14: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.2913202,
final_local_scores=1.2995095, pem=-4.928425,
similarity=69.67121]
15: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.7186224,
final_local_scores=0.6799845, pem=-3.8567436,
similarity=45.355377]
16: 22523 Oakland_Athletics
```

```
[pem_pos=7, final_scores=1.317741,
final_local_scores=1.3250294, pem=-3.9188845,
similarity=64.94273]
17: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.179795,
final_local_scores=1.1726778, pem=-4.191953,
similarity=61.881252]
18: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.0352947,
final_local_scores=1.0159657, pem=-4.1401463,
similarity=56.92292]
19: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.52170104,
final_local_scores=0.45940703, pem=-3.40205,
similarity=36.300213]
20: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.2628081,
final_local_scores=1.2650235, pem=-3.9188845,
similarity=63.15204]
21: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.0743716,
final_local_scores=1.0581775, pem=-4.046561,
similarity=57.673668]
22: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=1.2113485,
final_local_scores=1.2103457, pem=-4.2906566,
similarity=63.5421]
23: 64646 Toronto
    [pem_pos=0, final_scores=0.30088168,
final_local_scores=0.35240713, pem=-0.12291032,
similarity=15.274651]
24: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.1904346,
final_local_scores=1.1893072, pem=-4.928425,
similarity=66.38255]
25: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.83564764,
final_local_scores=0.80265, pem=-3.5915844,
similarity=47.573982]
26: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.1290002,
final_local_scores=1.1214931, pem=-4.418931,
similarity=61.588142]
27: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.18450458,
final_local_scores=0.08810252, pem=-5.0692368,
similarity=34.286167]
28: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.72706133,
final_local_scores=0.6892027, pem=-3.8567436,
similarity=45.630466]
29: 2456294 National_League_East
```

```
[pem_pos=0, final_scores=0.43273938,
final_local_scores=0.4264409, pem=0.0, similarity=16.81556]
30: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=1.1648302,
final_local_scores=1.1567581, pem=-4.136474,
similarity=61.10447]
31: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.7150133,
final_local_scores=0.67228305, pem=-4.5098605,
similarity=48.677303]
32: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.5637906,
final_local_scores=0.5023554, pem=-5.0692368,
similarity=46.648296]
33: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.2619226,
final_local_scores=1.2643852, pem=-4.5189333,
similarity=66.39615]
34: 2456377 National_League_Central
    [pem_pos=0, final_scores=0.5779788,
final_local_scores=0.5966241, pem=-0.98951703,
similarity=27.275314]
35: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.3473358,
final_local_scores=1.3583508, pem=-4.650927,
similarity=69.918076]
36: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.106658,
final_local_scores=1.094681, pem=-3.771247,
similarity=57.26581]
37: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.3115056,
final_local_scores=1.3165791, pem=-4.03702, similarity=65.333]
38: 2456474 National_League_West
    [pem_pos=1, final_scores=0.5820421,
final_local_scores=0.5887663, pem=-1.2952092,
similarity=28.703222]
39: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.938701,
final_local_scores=0.91660243, pem=-3.9307637,
similarity=52.81906]
40: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.2153736,
final_local_scores=1.2217512, pem=-4.482383,
similarity=64.9251]
41: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.93231374,
final_local_scores=0.90297025, pem=-4.409061,
similarity=55.013306]
42: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.0203698,
final_local_scores=1.0004244, pem=-3.771247,
```

```
similarity=54.452995]
43: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.1732115,
final_local_scores=1.1756954, pem=-4.482383,
similarity=63.5507]
44: 7954681 Montreal
    [pem_pos=0, final_scores=0.32627457,
final_local_scores=0.4008562, pem=-0.11023275,
similarity=16.651524]
45: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=1.0154637,
final_local_scores=0.9935979, pem=-4.136474,
similarity=56.235447]
46: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.2073419,
final_local_scores=1.2027962, pem=-4.03702,
similarity=61.937496]
47: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.97425205,
final_local_scores=0.9554365, pem=-3.9307637,
similarity=53.977947]
48: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.34763283,
final_local_scores=0.26623577, pem=-5.0692368,
similarity=39.602016]
49: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.4635109,
final_local_scores=1.4852543, pem=-4.650927,
similarity=73.70513]
50: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.2331648,
final_local_scores=1.2329717, pem=-4.5189333,
similarity=65.45871]
51: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.78345805,
final_local_scores=0.74036807, pem=-4.409061,
similarity=50.16093]
1: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.242411,
final_local_scores=0.15135652, pem=-5.0692368,
similarity=36.17379]
2: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.27590877,
final_local_scores=0.19582665, pem=-4.422637,
similarity=33.98456]
```

Sample: b'665_BASEBALL'

```
BASEBALL - [1 1 MAJOR LEAGUE 1] RESULTS TUESDAY . [1 1 NEW YO
RK 1] 1996-08-28 Results of [2 2 Major League 2] Baseball
games played on Tuesday ( home team in CAPS ) : [3 3 National
League 3] [4 4 Philadelphia 4] 3 [5 5 SAN FRANCISCO 5] 2 [6 6
Los Angeles 6] 5 [7 7 MONTREAL 7] 1 [8 8 PITTSBURGH 8] 3 [9 9
```

Atlanta 9] 2 [10 10 San Diego 10] 4 [2 2 NEW YORK 2] 3 [11 11
HOUSTON 11] 6 3 Chicago 5 4 Florida 6 ST LOUIS 3 [12 12 Cinci
nnati 12] 4 [13 13 COLORADO 13] 3 [14 14 American League 14] [
15 15 Cleveland 15] 12 [16 16 DETROIT 16] 2 [17 17 BALTIMORE 1
7] 3 [18 18 Oakland 18] 1 [19 19 Minnesota 19] 6 [20 20 TORON
TO 20] 4 [21 21 Milwaukee 21] 4 5 CHICAGO 2 [22 22 KANSAS CITY
22] 4 [23 23 Texas 23] 3 (10 innings) [24 24 Boston 24] 2 6
CALIFORNIA 1 [25 25 SEATTLE 25] 7 [26 26 New York 26] 4

1: 38776 Major_League_Baseball
2: 38776 Major_League_Baseball
3: 64290 National_League
4: 23741 Philadelphia_Phillies
5: 28416 San_Francisco_Giants
6: 18213 Los_Angeles_Dodgers
7: 7954681 Montreal
8: 24270 Pittsburgh_Pirates
9: 2140 Atlanta_Braves
10: 28417 San_Diego_Padres
11: 13894 Houston_Astros
12: 6671 Cincinnati_Reds
13: 6669 Colorado_Rockies
14: 64777 American_League
15: 6652 Cleveland_Indians
16: 8579 Detroit_Tigers
17: 4340 Baltimore_Orioles
18: 22523 Oakland_Athletics
19: 20050 Minnesota_Twins
20: 64646 Toronto
21: 20152 Milwaukee_Brewers
22: 16883 Kansas_City_Royals
23: 30857 Texas_Rangers_(baseball)
24: 4339 Boston_Red_Sox
25: 27660 Seattle_Mariners
26: 4848143 New_York_Yankees

1: 8210131 New_York
[winner: 645042 New_York_City, pem_pos=1,
final_scores=0.26185352, final_local_scores=0.3006919,
pem=-1.7906687, similarity=22.800894]
gt: 8210131 New_York, pem_pos=0,
final_scores=-0.026956096, final_local_scores=-0.007257521,
pem=-0.2889045, similarity=5.4442415]

2: 645042 New_York_City
[winner: 4848143 New_York_Yankees, pem_pos=11,
final_scores=0.2381122, final_local_scores=0.14736277,
pem=-5.0692368, similarity=36.05461]
gt: 645042 New_York_City, pem_pos=1,
final_scores=0.073962316, final_local_scores=0.095255196,
pem=-1.7906687, similarity=16.670252]

3: 5945 Chicago_White_Sox
[winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.69277936, final_local_scores=-0.78104705,
pem=-4.591283, similarity=5.7498198]

RECALL MISS: 5945 Chicago_White_Sox]
4: 11278 Miami_Marlins
[winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.7911695, final_local_scores=-0.86472744,
pem=-4.620281, similarity=3.410328]
RECALL MISS: 11278 Miami_Marlins]
5: 5945 Chicago_White_Sox
[winner: 6886 Chicago, pem_pos=0,
final_scores=-0.7211707, final_local_scores=-0.80264074,
pem=-0.20658566, similarity=-18.739243]
RECALL MISS: 5945 Chicago_White_Sox]
6: 1360083 Los_Angeles_Angels_ofAnaheim
[winner: 5407 California, pem_pos=0,
final_scores=-0.6306973, final_local_scores=-0.65989095,
pem=-0.08662026, similarity=-15.131695]
RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
1: 38776 Major_League_Baseball
[pem_pos=0, final_scores=0.5488687,
final_local_scores=0.49435943, pem=-0.47792417,
similarity=21.44141]
2: 38776 Major_League_Baseball
[pem_pos=0, final_scores=0.56088704,
final_local_scores=0.5074876, pem=-0.47792417,
similarity=21.83318]
3: 64290 National_League
[pem_pos=0, final_scores=0.5949766,
final_local_scores=0.5646797, pem=-0.07431644,
similarity=21.345026]
4: 23741 Philadelphia_Phillies
[pem_pos=10, final_scores=1.0327178,
final_local_scores=1.0160971, pem=-4.5189333,
similarity=58.986736]
5: 28416 San_Francisco_Giants
[pem_pos=3, final_scores=0.86883813,
final_local_scores=0.8348265, pem=-4.409061,
similarity=52.97976]
6: 18213 Los_Angeles_Dodgers
[pem_pos=15, final_scores=0.63596314,
final_local_scores=0.58072394, pem=-4.7524934,
similarity=47.264473]
7: 7954681 Montreal
[pem_pos=0, final_scores=0.36223906,
final_local_scores=0.43139702, pem=-0.11023275,
similarity=17.562923]
8: 24270 Pittsburgh_Pirates
[pem_pos=7, final_scores=1.1768152,
final_local_scores=1.1709249, pem=-4.03702,
similarity=60.98639]
9: 2140 Atlanta_Braves
[pem_pos=7, final_scores=0.88540083,
final_local_scores=0.8535755, pem=-4.136474,
similarity=52.0569]

```
10: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.54038984,
final_local_scores=0.48523724, pem=-3.9307637,
similarity=39.946262]
11: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.2713751,
final_local_scores=1.278784, pem=-4.650927,
similarity=67.54365]
12: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.0882151,
final_local_scores=1.0766404, pem=-3.771247,
similarity=56.727436]
13: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.0281565,
final_local_scores=1.0218961, pem=-4.482383,
similarity=58.96103]
14: 64777 American_League
    [pem_pos=0, final_scores=0.5423712,
final_local_scores=0.51267207, pem=0.0, similarity=19.38887]
15: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.0472624,
final_local_scores=1.0293217, pem=-4.191953,
similarity=57.603214]
16: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.142668,
final_local_scores=1.1349108, pem=-4.1401463,
similarity=60.472473]
17: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.2422584,
final_local_scores=1.2432284, pem=-4.046561,
similarity=63.195957]
18: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.249974,
final_local_scores=1.25336, pem=-3.9188845,
similarity=62.80398]
19: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=0.94472295,
final_local_scores=0.9203573, pem=-4.2906566,
similarity=54.888268]
20: 64646 Toronto
    [pem_pos=0, final_scores=0.39005852,
final_local_scores=0.4402631, pem=-0.12291032,
similarity=17.896448]
21: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.5437658,
final_local_scores=0.4870072, pem=-3.40205,
similarity=37.123856]
22: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.8614109,
final_local_scores=0.83288074, pem=-3.5915844,
similarity=48.47613]
23: 30857 Texas_Rangers_(baseball)
```

```
[pem_pos=13, final_scores=1.2425444,
final_local_scores=1.2489679, pem=-4.928425,
similarity=68.16295]
24: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.149296,
final_local_scores=1.1450768, pem=-4.418931,
similarity=62.291924]
25: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.744196,
final_local_scores=0.7104922, pem=-3.8567436,
similarity=46.26579]
26: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.5291886,
final_local_scores=0.46567482, pem=-5.0692368,
similarity=45.553677]
1: 645042 New_York_City
    [pem_pos=1, final_scores=0.26185352,
final_local_scores=0.3006919, pem=-1.7906687,
similarity=22.800894]
2: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.2381122,
final_local_scores=0.14736277, pem=-5.0692368,
similarity=36.05461]
```

Sample: b'666_TENNIS'

TENNIS - [1 1 GRAF 1] WORKS HARD FOR FIRST-ROUND WIN . Bill Berkrot [2 2 NEW YORK 2] 1996-08-27 It was n't supposed to be this hard for defending champion [3 3 Steffi Graf 3] to win her opening match at the [4 4 U.S. Open 4] on Tuesday night . But the script that called for the usual first-round demolition by the top-ranked top seed was rewritten by 29th-ranked [5 5 Indonesian 5] [6 6 Yayuk Basuki 6] playing with nothing to lose abandon . [7 7 Graf 7] , of course , prevailed 6-3 7-6 , but not before some tense moments that even had the [8 8 German 8] superstar thinking the match was going three sets . " I won the second set , which I did n't think I would do , being down 5-2 and the chances she had at 6-5 , " [9 9 Graf 9] recalled . Several of the other women 's seeds eased into the second round with more typical Graf-like efficiency Tuesday . As afternoon turned to evening , fourth-seeded [10 10 Spaniard 10] [11 11 Conchita Martinez 11] took apart [12 12 Romanian 12] [13 13 Ruxandra Dragomir 13] in 58 minutes with the loss of just two games , one more than second seed [14 14 Monica Seles 14] , who opened the second-day programme by crushing [15 15 American 15] Anne Miller 6-0 6-1 . Third seed [16 16 Arantxa Sanchez Vicario 16] , the 1994 champion , and eighth-seeded [17 17 Olympic 17] gold medalist [18 18 Lindsay Davenport 18] dropped three game each en route to the second round . But the day was not without its seeded casualties on the women 's side . Fifth-seed [19 19 Iva Majoli 19] of [20 20 Croatia 20] was picked off by [21 21 Austrian 21] [22 22 Judith Wiesner 22] and [23 23 Wimbledon 23]

] semifinalist [24 24 Kimiko Date 24] of [25 25 Japan 25] , the 10th seed , fell 6-2 7-5 to 53rd-ranked [26 26 American 26] [27 27 Kimberly Po 27] . Date 's defeat left no other seeded players in [28 28 Seles 28] 's quarter of the draw , which lost [29 29 Anke Huber 29] (6) and Maggie Maleeva (12) on Monday . But [30 30 Graf 30] , winner of 20 [31 31 Grand Slam 31] titles , was not about to join that list . " At some points I felt a little nervous , " she admitted . " When it came down to the important points , I felt more confident . " [32 32 Basuki 32] , a first-round loser here for the fifth consecutive year , was clearly going for winners , hitting the lines and running [33 33 Graf 33] around the court as she broke the top seed twice in the second set to grab that shocking 5-2 lead . [34 34 Graf 34] ran off the next three games to restore some semblance of order . But [35 35 Basuki 35] , her long black ponytail flying as she raced for shots , held her serve and twice had set point on [36 36 Graf 36] 's serve at 6-5 before the [37 37 German 37] unleashed a forehand pass to force the tie-break . " I lost the moment , " lamented [38 38 Basuki 38] , who has reached the fourth round at [39 39 Wimbledon 39] four times and was a semifinalist in [40 40 Montreal 40] earlier this month . Still , the feisty [41 41 Indonesian 41] got off to a 3-0 lead in the tie-breaker before a pair of costly double faults gave [42 42 Graf 42] her chance to avoid a third set . " Usually in the first one or two matches , you want to find your rhythm and want to get into it , " said [43 43 Graf 43] , who won seven of the last eight points in the breaker . " To be in that situation today , a couple of times having to play well , get down and play point-by-point , definitely is a good start . "

1: 164640 Steffi_Graf
2: 645042 New_York_City
3: 164640 Steffi_Graf
4: 82729 US_Open_(tennis)
5: 14579 Indonesia
6: 2672548 Yayuk_Basuki
7: 164640 Steffi_Graf
8: 11867 Germany
9: 164640 Steffi_Graf
10: 26667 Spain
11: 784152 Conchita_Martínez
12: 25445 Romania
13: 4038121 Ruxandra_Dragomir
14: 300468 Monica_Seles
15: 3434750 United_States
16: 534624 Arantxa_Sánchez_Vicario
17: 27566 Summer_Olympic_Games
18: 52828 Lindsay_Davenport
19: 928220 Iva_Majoli
20: 5573 Croatia
21: 26964606 Austria

```
22: 2805420 Judith_Wiesner
23: 91195 The_Championships,_Wimbledon
24: 495296 Kimiko_Date-Krumm
25: 15573 Japan
26: 3434750 United_States
27: 6426517 Kimberly_Po
28: 300468 Monica_Seles
29: 1443478 Anke_Huber
30: 164640 Steffi_Graf
31: 197638 Grand_Slam_(tennis)
32: 2672548 Yayuk_Basuki
33: 164640 Steffi_Graf
34: 164640 Steffi_Graf
35: 2672548 Yayuk_Basuki
36: 164640 Steffi_Graf
37: 11867 Germany
38: 2672548 Yayuk_Basuki
39: 91195 The_Championships,_Wimbledon
40: 7954681 Montreal
41: 14579 Indonesia
42: 164640 Steffi_Graf
43: 164640 Steffi_Graf
1: 164640 Steffi_Graf
    [pem_pos=0, final_scores=0.9011032,
final_local_scores=0.88420886, pem=0.0, similarity=30.476269]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.32633138,
final_local_scores=0.40535176, pem=-1.7906687,
similarity=25.924152]
3: 164640 Steffi_Graf
    [pem_pos=0, final_scores=0.90847427,
final_local_scores=0.8910281, pem=0.0, similarity=30.679768]
4: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.5353543,
final_local_scores=0.5238996, pem=-1.2370842,
similarity=26.451378]
5: 14579 Indonesia
    [pem_pos=1, final_scores=0.43258905,
final_local_scores=0.54852265, pem=-0.74208045,
similarity=24.494272]
6: 2672548 Yayuk_Basuki
    [pem_pos=0, final_scores=0.78950673,
final_local_scores=0.82731616, pem=0.0, similarity=28.778477]
7: 164640 Steffi_Graf
    [pem_pos=0, final_scores=1.1431215,
final_local_scores=1.1485772, pem=0.0, similarity=38.365547]
8: 11867 Germany
    [pem_pos=1, final_scores=0.7596739,
final_local_scores=0.86252636, pem=-1.0465823,
similarity=35.5207]
9: 164640 Steffi_Graf
    [pem_pos=0, final_scores=1.4762878,
```

```
final_local_scores=1.5125108, pem=0.0, similarity=49.22605]
10: 26667 Spain
    [pem_pos=1, final_scores=0.5255006,
final_local_scores=0.61721075, pem=-0.7552963,
similarity=26.61593]
11: 784152 Conchita_Martínez
    [pem_pos=0, final_scores=0.81099254,
final_local_scores=0.84052116, pem=0.0, similarity=29.17254]
12: 25445 Romania
    [pem_pos=1, final_scores=0.5903435,
final_local_scores=0.6822417, pem=-1.0382323,
similarity=30.095236]
13: 4038121 Ruxandra_Dragomir
    [pem_pos=0, final_scores=0.8644944,
final_local_scores=0.9501001, pem=0.0, similarity=32.442596]
14: 300468 Monica_Seles
    [pem_pos=0, final_scores=0.74197763,
final_local_scores=0.73152703, pem=0.0, similarity=25.919939]
15: 3434750 United_States
    [pem_pos=0, final_scores=0.66967994,
final_local_scores=0.7767681, pem=-0.12994486,
similarity=27.976683]
16: 534624 Arantxa_Sánchez_Vicario
    [pem_pos=0, final_scores=0.5856761,
final_local_scores=0.5818762, pem=0.0, similarity=21.45406]
17: 27566 Summer_Olympic_Games
    [pem_pos=2, final_scores=0.35327566,
final_local_scores=0.39525038, pem=-3.2532303,
similarity=33.576347]
18: 52828 Lindsay_Davenport
    [pem_pos=0, final_scores=0.89808637,
final_local_scores=0.9125239, pem=0.0, similarity=31.321247]
19: 928220 Iva_Majoli
    [pem_pos=0, final_scores=0.8898466,
final_local_scores=0.93931335, pem=0.0, similarity=32.120697]
20: 5573 Croatia
    [pem_pos=0, final_scores=0.5837715,
final_local_scores=0.6710675, pem=-0.35021794,
similarity=26.020245]
21: 26964606 Austria
    [pem_pos=0, final_scores=0.6193106,
final_local_scores=0.7055446, pem=-0.40674555,
similarity=27.356516]
22: 2805420 Judith_Wiesner
    [pem_pos=0, final_scores=0.64694065,
final_local_scores=0.73680794, pem=0.0, similarity=26.077532]
23: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.5217881,
final_local_scores=0.50528765, pem=-0.8505555,
similarity=23.793957]
24: 495296 Kimiko_Date-Krumm
    [pem_pos=0, final_scores=0.6829255,
```

```
final_local_scores=0.69877094, pem=0.0, similarity=24.942432]
25: 15573 Japan
    [pem_pos=0, final_scores=0.56107545,
final_local_scores=0.6720941, pem=-0.13549899,
similarity=24.883205]
26: 3434750 United_States
    [pem_pos=0, final_scores=0.58741355,
final_local_scores=0.68690455, pem=-0.12994486,
similarity=25.294975]
27: 6426517 Kimberly_Po
    [pem_pos=0, final_scores=0.67194206,
final_local_scores=0.7410967, pem=0.0, similarity=26.205517]
28: 300468 Monica_Seles
    [pem_pos=0, final_scores=0.59526473,
final_local_scores=0.5734562, pem=0.0, similarity=21.202791]
29: 1443478 Anke_Huber
    [pem_pos=0, final_scores=0.98244196,
final_local_scores=1.0424647, pem=0.0, similarity=35.198936]
30: 164640 Steffi_Graf
    [pem_pos=0, final_scores=1.3119484,
final_local_scores=1.3329949, pem=0.0, similarity=43.86894]
31: 197638 Grand_Slam_(tennis)
    [pem_pos=0, final_scores=0.5034726,
final_local_scores=0.48393947, pem=-0.37449613,
similarity=20.568]
32: 2672548 Yayuk_Basuki
    [pem_pos=0, final_scores=0.7580047,
final_local_scores=0.7967854, pem=0.0, similarity=27.867378]
33: 164640 Steffi_Graf
    [pem_pos=0, final_scores=1.3207985,
final_local_scores=1.3426623, pem=0.0, similarity=44.157433]
34: 164640 Steffi_Graf
    [pem_pos=0, final_scores=1.2847095,
final_local_scores=1.3032405, pem=0.0, similarity=42.981007]
35: 2672548 Yayuk_Basuki
    [pem_pos=0, final_scores=0.739468,
final_local_scores=0.77653676, pem=0.0, similarity=27.263119]
36: 164640 Steffi_Graf
    [pem_pos=0, final_scores=1.2599438,
final_local_scores=1.2761879, pem=0.0, similarity=42.173706]
37: 11867 Germany
    [pem_pos=1, final_scores=0.3616981,
final_local_scores=0.4277982, pem=-1.0465823,
similarity=22.547543]
38: 2672548 Yayuk_Basuki
    [pem_pos=0, final_scores=0.8974679,
final_local_scores=0.94912773, pem=0.0, similarity=32.413578]
39: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.55772275,
final_local_scores=0.54454076, pem=-0.8505555,
similarity=24.96535]
40: 7954681 Montreal
```

```
[pem_pos=0, final_scores=0.42528898,
final_local_scores=0.50567067, pem=-0.11023275,
similarity=19.779398]
41: 14579 Indonesia
    [pem_pos=1, final_scores=0.428073,
final_local_scores=0.5435896, pem=-0.74208045,
similarity=24.34706]
42: 164640 Steffi_Graf
    [pem_pos=0, final_scores=1.4643818,
final_local_scores=1.4995053, pem=0.0, similarity=48.837944]
43: 164640 Steffi_Graf
    [pem_pos=0, final_scores=1.502334,
final_local_scores=1.5409622, pem=0.0, similarity=50.0751]
```

Sample: b'667_BASEBALL'

BASEBALL - [1 1 SOSA 1] HAS SURGERY , OUT UP TO SIX WEEKS . [2 2 CHICAGO 2] 1996-08-27 [3 3 Chicago Cubs 3] right fielder [4 4 Sammy Sosa 4] underwent surgery on Monday to remove a fractured bone from his right hand and will miss four to six weeks , the club announced Tuesday . [5 5 Sosa 5] , a leading candidate for [6 6 National League 6] [7 7 Most Valuable Player 7] honours , was injured August 20th when he was hit by a [8 8 Mark Hutton 8] pitch in the first inning of an 8-1 victory over the [9 9 Florida Marlins 9] . The 27-year-old [10 10 Sosa 10] leads the league with 40 homers and is tied for 10th with 100 [11 11 RBI 11] . The loss of [12 12 Sosa 12] , who appeared in all 124 games this season , is a huge blow to the [13 13 Cubs 13] ' playoff hopes .

```
1: 26920 Sammy_Sosa
2: 6886 Chicago
3: 6654 Chicago_Cubs
4: 26920 Sammy_Sosa
5: 26920 Sammy_Sosa
6: 64290 National_League
7: 3858 Major_League_Baseball_Most_Valuable_Player_Award
8: 3526509 Mark_Hutton
9: 11278 Miami_Marlins
10: 26920 Sammy_Sosa
11: 300747 Run_(baseball)
12: 26920 Sammy_Sosa
13: 6654 Chicago_Cubs
1: 26920 Sammy_Sosa
    [pem_pos=0, final_scores=0.9539054,
final_local_scores=0.9415, pem=0.0, similarity=32.18595]
2: 6886 Chicago
    [pem_pos=0, final_scores=0.5010656,
final_local_scores=0.60560244, pem=-0.20658566,
similarity=23.285543]
3: 6654 Chicago_Cubs
    [pem_pos=0, final_scores=0.49549168,
final_local_scores=0.45944113, pem=-0.085107334,
similarity=18.26318]
```

```
4: 26920 Sammy_Sosa
    [pem_pos=0, final_scores=0.73230445,
final_local_scores=0.6961244, pem=0.0, similarity=24.863453]
5: 26920 Sammy_Sosa
    [pem_pos=0, final_scores=0.934177,
final_local_scores=0.9199497, pem=0.0, similarity=31.542849]
6: 64290 National_League
    [pem_pos=0, final_scores=0.4120149,
final_local_scores=0.38621497, pem=-0.07431644,
similarity=16.019283]
7: 3858 Major_League_Baseball_Most_Valueable_Player_Award
    [pem_pos=1, final_scores=0.46956807,
final_local_scores=0.44657642, pem=-1.3928809,
similarity=24.991148]
8: 3526509 Mark_Hutton
    [pem_pos=0, final_scores=0.476444842,
final_local_scores=0.5943528, pem=0.0, similarity=21.826385]
9: 11278 Miami_Marlins
    [pem_pos=0, final_scores=0.52104837,
final_local_scores=0.50350523, pem=0.0, similarity=19.115313]
10: 26920 Sammy_Sosa
    [pem_pos=0, final_scores=0.92896575,
final_local_scores=0.9142572, pem=0.0, similarity=31.372972]
11: 300747 Run_(baseball)
    [pem_pos=2, final_scores=0.57000214,
final_local_scores=0.56997544, pem=-1.0986123,
similarity=27.073341]
12: 26920 Sammy_Sosa
    [pem_pos=0, final_scores=0.96998066,
final_local_scores=0.9590598, pem=0.0, similarity=32.70997]
13: 6654 Chicago_Cubs
    [pem_pos=0, final_scores=0.49591845,
final_local_scores=0.45990735, pem=-0.2883288,
similarity=19.382242]
```

Sample: b'668_SOCCER'
SOCCER - [1 1 MARCELO 1] HAT-TRICK KEEPS [2 2 PSV 2] AT TOP
OF [3 3 DUTCH 3] LEAGUE . [4 4 AMSTERDAM 4] 1996-08-28 [5 5 Br
azilian 5] striker [6 6 Marcelo 6] scored a hat-trick as [7 7
PSV Eindhoven 7] maintained their 100 percent record and
stayed on top of the [8 8 Dutch 8] first division with a 3-1
win at [9 9 Volendam 9] on Wednesday . [10 10 PSV 10] 's main
rivals for the title , defending champions [11 11 Ajax Amster
dam 11] , celebrated the novelty of having the roof of their
new 51,000 seat stadium closed against the rain , with a 1-0
win over [12 12 AZ Alkmaar 12] . [13 13 Ajax 13] were missing
six first-team players but [14 14 Frank de Boer 14] shot home
the winner from a 20-metre free kick in the 30th minute of a
dull game . [15 15 Marcelo 15] , signed in close season to
replace compatriot [16 16 Ronaldo 16] who left to play for [17
17 Barcelona 17] , opened the [18 18 PSV 18] scoring in the
19th minute when he fired home after good work from [19 19 Re

ne Eijkelkamp 19] . The [20 20 Brazilian 20] found the mark again two minutes after halftime and again in the 56th minute before midfielder Pascal Jongsma scored a consolation goal for [21 21 Volendam 21] five minutes from time . [22 22 Feyenoord Rotterdam 22] suffered an early shock when they went 1-0 down after four minutes against de Graafschap Doetinchem . The equaliser came in the 73rd minute when [23 23 Swedish 23] international [24 24 Henke Larsson 24] scored from close range and 10 minutes later [25 25 Jean-Paul van Gastel 25] gave [26 26 Feyenoord 26] a 2-1 victory from the penalty spot . After three matches [27 27 PSV 27] lead the first division with nine points , three points clear of fifth-placed [28 28 Ajax 28] .

1: 16001137 Marcelo_Silva_Ramos
2: 79820 PSV_Eindhoven
3: 21148 Netherlands
4: 844 Amsterdam
5: 3383 Brazil
6: 16001137 Marcelo_Silva_Ramos
7: 79820 PSV_Eindhoven
8: 21148 Netherlands
9: 2202202 FC_Volendam
10: 79820 PSV_Eindhoven
11: 2273 AFC_Ajax
12: 2289 AZ_Alkmaar
13: 2273 AFC_Ajax
14: 685945 Frank_de_Boer
15: 16001137 Marcelo_Silva_Ramos
16: 66164 Ronaldo
17: 68187 FC_Barcelona
18: 79820 PSV_Eindhoven
19: 9464508 René_Eijkelkamp
20: 3383 Brazil
21: 2202202 FC_Volendam
22: 47707 Feyenoord
23: 5058739 Sweden
24: 240952 Henrik_Larsson
25: 9420542 Jean-Paul_van_Gastel
26: 47707 Feyenoord
27: 79820 PSV_Eindhoven
28: 2273 AFC_Ajax
1: 16001137 Marcelo_Silva_Ramos
 [pem_pos=13, final_scores=0.09365426,
final_local_scores=0.18735152, pem=-3.61824,
similarity=29.357208]
2: 79820 PSV_Eindhoven
 [pem_pos=0, final_scores=0.5601962,
final_local_scores=0.50883394, pem=0.0, similarity=19.274332]
3: 21148 Netherlands
 [pem_pos=0, final_scores=0.5628354,
final_local_scores=0.5929071, pem=-0.7943452,
similarity=26.103016]

4: 844 Amsterdam
[pem_pos=0, final_scores=0.5542751,
final_local_scores=0.5930488, pem=-0.0642803,
similarity=22.13704]
5: 3383 Brazil
[pem_pos=0, final_scores=0.6684757,
final_local_scores=0.7534453, pem=-0.2733083,
similarity=28.060316]
6: 16001137 Marcelo_Silva_Ramos
[pem_pos=13, final_scores=0.6167881,
final_local_scores=0.75718087, pem=-3.61824,
similarity=46.362053]
7: 79820 PSV_Eindhoven
[pem_pos=0, final_scores=0.56754404,
final_local_scores=0.5168603, pem=0.0, similarity=19.513855]
8: 21148 Netherlands
[pem_pos=0, final_scores=0.5066089,
final_local_scores=0.5314882, pem=-0.7943452,
similarity=24.270153]
9: 2202202 FC_Volendam
[pem_pos=1, final_scores=0.50189024,
final_local_scores=0.48894322, pem=-0.8958714,
similarity=23.55264]
10: 79820 PSV_Eindhoven
[pem_pos=0, final_scores=0.55625635,
final_local_scores=0.50453025, pem=0.0, similarity=19.145903]
11: 2273 AFC_Ajax
[pem_pos=0, final_scores=0.56628054,
final_local_scores=0.51680905, pem=0.0, similarity=19.512325]
12: 2289 AZ_Alkmaar
[pem_pos=0, final_scores=0.44050103,
final_local_scores=0.39596027, pem=0.0, similarity=15.905957]
13: 2273 AFC_Ajax
[pem_pos=1, final_scores=0.5569379,
final_local_scores=0.50660354, pem=-1.4356393,
similarity=27.015003]
14: 685945 Frank_de_Boer
[pem_pos=0, final_scores=0.6700879,
final_local_scores=0.70238346, pem=0.0, similarity=25.050236]
15: 16001137 Marcelo_Silva_Ramos
[pem_pos=13, final_scores=0.58246654,
final_local_scores=0.7196898, pem=-3.61824,
similarity=45.243244]
16: 66164 Ronaldo
[pem_pos=0, final_scores=0.31398958,
final_local_scores=0.31589675, pem=-0.10865249,
similarity=14.107571]
17: 68187 FC_Barcelona
[pem_pos=1, final_scores=0.44636315,
final_local_scores=0.42223305, pem=-1.7089969,
similarity=25.983778]
18: 79820 PSV_Eindhoven

```
[pem_pos=0, final_scores=0.4935041,
final_local_scores=0.43598294, pem=0.0, similarity=17.100313]
19: 9464508 René_Eijkelkamp
    [pem_pos=0, final_scores=0.49796665,
final_local_scores=0.59538037, pem=0.0, similarity=21.85705]
20: 3383 Brazil
    [pem_pos=0, final_scores=0.6440131,
final_local_scores=0.72672373, pem=-0.2733083,
similarity=27.262892]
21: 2202202 FC_Volendam
    [pem_pos=1, final_scores=0.6476044,
final_local_scores=0.6481138, pem=-0.8958714,
similarity=28.302608]
22: 47707 Feyenoord
    [pem_pos=0, final_scores=0.57249486,
final_local_scores=0.53130543, pem=0.0, similarity=19.944927]
23: 5058739 Sweden
    [pem_pos=0, final_scores=0.71041006,
final_local_scores=0.7902188, pem=-0.63135606,
similarity=31.104832]
24: 240952 Henrik_Larsson
    [pem_pos=0, final_scores=0.6787563,
final_local_scores=0.7217585, pem=0.0, similarity=25.628426]
25: 9420542 Jean-Paul_van_Gastel
    [pem_pos=0, final_scores=0.6071088,
final_local_scores=0.6962213, pem=0.0, similarity=24.866344]
26: 47707 Feyenoord
    [pem_pos=0, final_scores=0.5350506,
final_local_scores=0.49040318, pem=0.0, similarity=18.724321]
27: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.58283335,
final_local_scores=0.5335616, pem=0.0, similarity=20.012255]
28: 2273 AFC_Ajax
    [pem_pos=1, final_scores=0.51572007,
final_local_scores=0.46157938, pem=-1.4356393,
similarity=25.67139]
```

Sample: b'669_SOCCER'

SOCCER - [1 1 ROBSON 1] WINS FIRST TROPHY WITH [2 2 BARCELONA 2] . [3 3 MADRID 3] 1996-08-28 Former [4 4 England 4] manager [5 5 Bobby Robson 5] enjoyed his first success in charge of [6 6 Barcelona 6] as his team weathered 90 minutes of non-stop [7 7 Atletico Madrid 7] pressure to win the [8 8 Spanish 8] Super Cup 6-5 on aggregate on Wednesday . [9 9 Barcelona 9] had won the first leg 5-2 but the second leg was a different story . Atletico came within a whisker of taking the Cup on the away-goal rule but squandered several chances after going 3-1 ahead 15 minutes from the end . Juan Lopez gave Atletico the lead midway through the first half after [10 10 Barcelona 10] fullback [11 11 Albert Ferrer 11] and substitute goalkeeper [12 12 Julen Lopetegui 12] failed to clear a [13 13 Milinko Pantic 13] cross . [14 14 Barcelona 14] 's [15 15 Hri

sto Stoichkov 15] made his only significant contribution of the evening 10 minutes after halftime when [16 16 Sergi Barjuan 16] broke down the right to set up the fiery [17 17 Bulgarian 17] with a simple equaliser . But Atletico struck back almost immediately through new signing [18 18 Juan Eduardo Esnaider 18] and then [19 19 Serbian 19] set-piece specialist Pantic made it 3-1 with a superb free-kick in the 75th minute . [20 20 Robson 20] praised Atletico after the game , which was played in the [21 21 Community of Madrid 21] athletic stadium because of pitch problems at the Vicente Calderon ground . The venue of Atletico 's first league game , scheduled for Sunday , is still in doubt with the [22 22 Real Madrid 22] 's [23 23 Santiago Bernabeu 23] a distinct possibility .

1: 68685 Bobby_Robson
2: 68187 FC_Barcelona
3: 41188263 Madrid
4: 9904 England_national_football_team
5: 68685 Bobby_Robson
6: 68187 FC_Barcelona
7: 84197 Atlético_Madrid
8: 26667 Spain
9: 68187 FC_Barcelona
10: 68187 FC_Barcelona
11: 3250276 Albert_Ferrer
12: 9279674 Julen_Lopetegui
13: 5431085 Milinko_Pantić
14: 68187 FC_Barcelona
15: 28714631 Hristo_Stoichkov
16: 1786727 Sergi_Barjuán
17: 3415 Bulgaria
18: 8146787 Juan_Esnáider
19: 29265 Serbia
20: 68685 Bobby_Robson
21: 69802 Community_of_Madrid
22: 26413 Real_Madrid_C.F.
23: 225760 Santiago_Bernabéu_Stadium
1: 68685 Bobby_Robson
 [pem_pos=0, final_scores=0.71853113,
final_local_scores=0.7354969, pem=0.0, similarity=26.038406]
2: 68187 FC_Barcelona
 [pem_pos=1, final_scores=0.5271135,
final_local_scores=0.47013444, pem=-1.7089969,
similarity=27.413252]
3: 41188263 Madrid
 [pem_pos=0, final_scores=0.49251658,
final_local_scores=0.5326997, pem=-0.13236846,
similarity=20.706375]
4: 9904 England_national_football_team
 [pem_pos=1, final_scores=0.41270453,
final_local_scores=0.3887378, pem=-2.776464,
similarity=30.789267]

5: 68685 Bobby_Robson
[pem_pos=0, final_scores=0.55770904,
final_local_scores=0.5493343, pem=0.0, similarity=20.482944]
6: 68187 FC_Barcelona
[pem_pos=1, final_scores=0.54583913,
final_local_scores=0.49058932, pem=-1.7089969,
similarity=28.023666]
7: 84197 Atlético_Madrid
[pem_pos=0, final_scores=0.4319058,
final_local_scores=0.3880884, pem=0.0, similarity=15.671045]
8: 26667 Spain
[pem_pos=1, final_scores=0.64288193,
final_local_scores=0.7100931, pem=-1.2425895,
similarity=32.0377]
9: 68187 FC_Barcelona
[pem_pos=1, final_scores=0.5012328,
final_local_scores=0.44186372, pem=-1.7089969,
similarity=26.5696]
10: 68187 FC_Barcelona
[pem_pos=1, final_scores=0.54100746,
final_local_scores=0.48531145, pem=-1.7089969,
similarity=27.866163]
11: 3250276 Albert_Ferrer
[pem_pos=0, final_scores=0.8207825,
final_local_scores=0.8730529, pem=0.0, similarity=30.143353]
12: 9279674 Julen_Lopetegui
[pem_pos=0, final_scores=0.6765061,
final_local_scores=0.732531, pem=0.0, similarity=25.9499]
13: 5431085 Milinko_Pantić
[pem_pos=0, final_scores=0.6389993,
final_local_scores=0.74245137, pem=0.0, similarity=26.245943]
14: 68187 FC_Barcelona
[pem_pos=1, final_scores=0.45114028,
final_local_scores=0.38714522, pem=-1.7089969,
similarity=24.936687]
15: 28714631 Hristo_Stoichkov
[pem_pos=0, final_scores=0.81557816,
final_local_scores=0.8530189, pem=0.0, similarity=29.545498]
16: 1786727 Sergi_Barjuán
[pem_pos=0, final_scores=0.66887707,
final_local_scores=0.724443, pem=0.0, similarity=25.708538]
17: 3415 Bulgaria
[pem_pos=1, final_scores=0.49070287,
final_local_scores=0.5864827, pem=-1.1560022,
similarity=27.878046]
18: 8146787 Juan_Esnáider
[pem_pos=0, final_scores=0.596361,
final_local_scores=0.620433, pem=0.0, similarity=22.60467]
19: 29265 Serbia
[pem_pos=1, final_scores=0.3279791,
final_local_scores=0.3949231, pem=-1.2504034,
similarity=22.674894]

```
20: 68685 Bobby_Robson
    [pem_pos=0, final_scores=0.7269232,
final_local_scores=0.74466395, pem=0.0, similarity=26.31197]
21: 69802 Community_of_Madrid
    [pem_pos=0, final_scores=0.37994063,
final_local_scores=0.41767925, pem=0.0, similarity=16.554094]
22: 26413 Real_Madrid_C.F.
    [pem_pos=0, final_scores=0.35908782,
final_local_scores=0.2914264, pem=0.0, similarity=12.786458]
23: 225760 Santiago_Bernabéu_Stadium
    [pem_pos=0, final_scores=0.37464422,
final_local_scores=0.34874728, pem=0.0, similarity=14.497027]
```

Sample: b'670_SOCCER'
SOCCER - SUMMARY OF [1 1 SPANISH 1] SUPER CUP . [2 2 MADRID 2] 1996-08-28 Summary of the [3 3 Spanish 3] Super Cup , second leg , played on Wednesday : [4 4 Atletico Madrid 4] 3 (Juan Lopez 28th minute , [5 5 Juan Esnaider 5] 58th , [6 6 Milinko Pantic 6] 75th) [7 7 Barcelona 7] 1 ([8 8 Hristo Stoichkov 8] 55th) . Halftime 1-0 . Attendance 11,000 . ([9 9 Barcelona 9] win 6-5 on aggregate) .

```
1: 26667 Spain
2: 41188263 Madrid
3: 26667 Spain
4: 84197 Atlético_Madrid
5: 8146787 Juan_Esnáider
6: 5431085 Milinko_Pantić
7: 68187 FC_Barcelona
8: 28714631 Hristo_Stoichkov
9: 68187 FC_Barcelona
1: 26667 Spain
    [pem_pos=2, final_scores=0.74144536,
final_local_scores=0.78559846, pem=-1.2425895,
similarity=34.29093]
2: 41188263 Madrid
    [pem_pos=0, final_scores=0.56601936,
final_local_scores=0.57351536, pem=-0.13236846,
similarity=21.924395]
3: 26667 Spain
    [pem_pos=1, final_scores=0.6915273,
final_local_scores=0.70797306, pem=-1.2425895,
similarity=31.974436]
4: 84197 Atlético_Madrid
    [pem_pos=0, final_scores=0.40649772,
final_local_scores=0.3705281, pem=0.0, similarity=15.147011]
5: 8146787 Juan_Esnáider
    [pem_pos=0, final_scores=0.60703164,
final_local_scores=0.6519719, pem=0.0, similarity=23.545855]
6: 5431085 Milinko_Pantić
    [pem_pos=0, final_scores=0.79249674,
final_local_scores=0.91268796, pem=0.0, similarity=31.326141]
7: 68187 FC_Barcelona
```

```
[pem_pos=1, final_scores=0.4486931,
final_local_scores=0.4061728, pem=-1.7089969,
similarity=25.504509]
8: 28714631 Hristo_Stoichkov
    [pem_pos=0, final_scores=0.6984608,
final_local_scores=0.7515678, pem=0.0, similarity=26.517994]
9: 68187 FC_Barcelona
    [pem_pos=1, final_scores=0.5402586,
final_local_scores=0.5061942, pem=-1.7089969,
similarity=28.489347]

Sample: b'671_SOCCER'
SOCCER - [1 1 BARCELONA 1] WIN [2 2 SPANISH 2] SUPER CUP . [3
3 MADRID 3] 1996-08-28 Result of the [4 4 Spanish 4] Super
Cup , second leg , played on Wednesday : [5 5 Atletico Madrid
5] 3 [6 6 Barcelona 6] 1 ( [7 7 Barcelona 7] win 6-5 on
aggregate )
1: 68187 FC_Barcelona
2: 26667 Spain
3: 41188263 Madrid
4: 26667 Spain
5: 84197 Atlético_Madrid
6: 68187 FC_Barcelona
7: 68187 FC_Barcelona
1: 68187 FC_Barcelona
    [pem_pos=1, final_scores=0.55595964,
final_local_scores=0.52697796, pem=-1.7089969,
similarity=29.109575]
2: 26667 Spain
    [pem_pos=2, final_scores=0.60617644,
final_local_scores=0.6192605, pem=-1.2425895,
similarity=29.327076]
3: 41188263 Madrid
    [pem_pos=0, final_scores=0.5666873,
final_local_scores=0.56030405, pem=-0.13236846,
similarity=21.530144]
4: 26667 Spain
    [pem_pos=1, final_scores=0.7044037,
final_local_scores=0.7037392, pem=-1.2425895,
similarity=31.84809]
5: 84197 Atlético_Madrid
    [pem_pos=0, final_scores=0.44864208,
final_local_scores=0.4307893, pem=0.0, similarity=16.945324]
6: 68187 FC_Barcelona
    [pem_pos=1, final_scores=0.5079858,
final_local_scores=0.4745738, pem=-1.7089969,
similarity=27.545732]
7: 68187 FC_Barcelona
    [pem_pos=1, final_scores=0.52119917,
final_local_scores=0.4890074, pem=-1.7089969,
similarity=27.976461]
```

Sample: b'672_SOCCER'
SOCCER - 1 AJAX SIGN [1 1 ARGENTINE 1] STRIKER GABRICH . [2 2
AMSTERDAM 2] 1996-08-28 [3 3 Argentine 3] striker Iwan Cesar
Gabrich signed a five year contract with [4 4 Dutch 4]
champions [5 5 Ajax Amsterdam 5] on Wednesday . The 24-year-
old Gabrich , who signed for an undisclosed fee from the [6 6
Argentine 6] side Newell Old Boys , is set to join [7 7 Dutch
7] international [8 8 Patrick Kluivert 8] in the [9 9 Ajax 9]
forward line . He is [10 10 Ajax 10] 's sixth new signing
this year , joining midfielder [11 11 Richard Witschge 11] ,
defenders [12 12 John Veldman 12] and [13 13 Mariano Juan 13]
and strikers [14 14 Tijjani Babangida 14] and [15 15 Dani 15]

.
1: 18951905 Argentina
2: 844 Amsterdam
3: 18951905 Argentina
4: 21148 Netherlands
5: 2273 AFC_Ajax
6: 18951905 Argentina
7: 21148 Netherlands
8: 615769 Patrick_Kluivert
9: 2273 AFC_Ajax
10: 2273 AFC_Ajax
11: 5823946 Richard_Witschge
12: 7363720 John_Veldman
13: 19834644 Mariano_Juan
14: 2803637 Tijjani_Babangida
15: 3521904 Daniel_da_Cruz_Carvalho
1: 2273 AFC_Ajax
[winner: 18539770 Jong_Ajax, pem_pos=14,
final_scores=-0.3251897, final_local_scores=-0.3635729,
pem=-4.0073333, similarity=15.032474]
RECALL MISS: 2273 AFC_Ajax]
1: 18951905 Argentina
[pem_pos=0, final_scores=0.43026787,
final_local_scores=0.44843453, pem=-0.18885723,
similarity=18.49893]
2: 844 Amsterdam
[pem_pos=0, final_scores=0.5677518,
final_local_scores=0.58162063, pem=-0.0642803,
similarity=21.796]
3: 18951905 Argentina
[pem_pos=0, final_scores=0.52801734,
final_local_scores=0.555211, pem=-0.18885723,
similarity=21.685352]
4: 21148 Netherlands
[pem_pos=0, final_scores=0.3712532,
final_local_scores=0.35069352, pem=-0.7943452,
similarity=18.87488]
5: 2273 AFC_Ajax
[pem_pos=0, final_scores=0.53578013,
final_local_scores=0.50816995, pem=0.0, similarity=19.254519]

```
6: 18951905 Argentina
    [pem_pos=0, final_scores=0.61188155,
final_local_scores=0.6468199, pem=-0.18885723,
similarity=24.419144]
7: 21148 Netherlands
    [pem_pos=0, final_scores=0.6654183,
final_local_scores=0.67202425, pem=-0.7943452,
similarity=28.46403]
8: 615769 Patrick_Kluivert
    [pem_pos=0, final_scores=0.8482664,
final_local_scores=0.92253596, pem=0.0, similarity=31.620026]
9: 2273 AFC_Ajax
    [pem_pos=1, final_scores=0.60631585,
final_local_scores=0.58521944, pem=-1.4356393,
similarity=29.361057]
10: 2273 AFC_Ajax
    [pem_pos=1, final_scores=0.57103544,
final_local_scores=0.546681, pem=-1.4356393,
similarity=28.210995]
11: 5823946 Richard_Witschge
    [pem_pos=0, final_scores=0.81702787,
final_local_scores=0.8945989, pem=0.0, similarity=30.786327]
12: 7363720 John_Veldman
    [pem_pos=0, final_scores=0.76080686,
final_local_scores=0.878244, pem=0.0, similarity=30.298265]
13: 19834644 Mariano_Juan
    [pem_pos=0, final_scores=0.40910953,
final_local_scores=0.5450475, pem=0.0, similarity=20.355019]
14: 2803637 Tijani_Babangida
    [pem_pos=0, final_scores=0.5795752,
final_local_scores=0.6366059, pem=0.0, similarity=23.087303]
15: 3521904 Daniel_da_Cruz_Carvalho
    [pem_pos=5, final_scores=0.6004431,
final_local_scores=0.6723458, pem=-2.8560238,
similarity=39.685352]
```

Sample: b'673_SOCCER'

SOCCER - [1 1 PARMA 1] , [2 2 ROMA 2] AND [3 3 UDINESE 3]
OUT OF [4 4 ITALIAN CUP 4] . [5 5 ROME 5] 1996-08-28 [1 1 UEFA
Cup 1] hopefuls [6 6 Parma 6] and [7 7 Roma 7] , under new
coaches this season , crashed out of the [8 8 Italian Cup 8]
to second division opponents on Wednesday while league
champions [9 9 Milan 9] could only draw 1-1 at humble [10 10 E
mpoli 10] . Wealthy [11 11 Parma 11] , now coached by the
former [12 12 Italian 12] international [13 13 Carlo Ancelotti
13] , were without new striker [14 14 Enrico Chiesa 14] and
went down 3-1 at serie B club Pescara in their second round
clash . [2 Pescara 2] 's Ottavio Palladini shattered [15 15 Pa
rma 15] with goals in the second and fourth minutes .
Midfielder Marco Giampaolo made it 3-0 in the 38th minute and
[16 16 Parma 16] 's [17 17 Alessandro Melli 17] pulled back a
late goal six minutes from time . The second round was the

entry point for the bulk of the [18 18 serie A 18] sides with the winners going through . The later stages of the cup are played over two legs . [19 19 Parma 19] 's defeat was a repeat of last season 's fiasco when they lost their opening cup match 3-0 to [20 20 Palermo 20] . [21 21 Roma 21] , now coached by [22 22 Argentine 22] [23 23 Carlos Bianchi 23] and watched by [24 24 Italian 24] national coach [25 25 Arrigo Sacci 25] , lost 3-1 to [26 26 Cesena 26] -- another repeat of last season when the [27 27 Rome 27] club also went out at the first hurdle . [28 28 Udinese 28] , with [29 29 Germany 29] 's Euro ' 96 hero [30 30 Oliver Bierhoff 30] in their lineup , completed the hat-trick of beaten [31 31 serie A 31] sides when they went under 2-1 to newly relegated [32 32 Cremonese 32] . [33 33 Milan 33] 's new [34 34 Uruguayan 34] coach [35 35 Oscar Tabarez 35] avoided the nightmare of defeat but faces a replay at home next Sunday . 2 Cup holders [36 36 Fiorentina 36] easily beat 3 Cosenza 3-1 while [37 37 European Cup 37] holders [38 38 Juventus 38] also cruised through with a 2-0 win at small southern club [39 39 Fidelis Andria 39] . Two other [40 40 serie A 40] sides lost at the weekend -- [41 41 Piacenza 41] and last year 's losing finalists [42 42 Atalanta 42] . Two cup matches could not be played on Wednesday due to argument over first round results . [43 43 Lecce 43] 's 3-0 weekend defeat of [44 44 Genoa 44] was expected to be overturned by a sporting judge on Thursday after the home club fielded an ineligible player . That would set [45 45 Genoa 45] up for a second round match against local rivals [46 46 Sampdoria 46] . [47 47 Nocerina 47] 's 4-3 defeat of [48 48 Piacenza 48] was also subject to a complaint , later removed , that forced their second round match against [49 49 serie A 49] newcomers [50 50 Perugia 50] to be delayed .

- 1: 62611 Parma_F.C.
- 2: 2358 A.S._Roma
- 3: 149877 Udinese_Calcio
- 4: 1281200 Coppa_Italia
- 5: 25458 Rome
- 6: 62611 Parma_F.C.
- 7: 2358 A.S._Roma
- 8: 1281200 Coppa_Italia
- 9: 18940588 A.C._Milan
- 10: 1136866 Empoli_F.C.
- 11: 62611 Parma_F.C.
- 12: 14532 Italy
- 13: 1142794 Carlo_Ancelotti
- 14: 1527955 Enrico_Chiesa
- 15: 62611 Parma_F.C.
- 16: 62611 Parma_F.C.
- 17: 18424280 Alessandro_Melli
- 18: 15496 Serie_A
- 19: 62611 Parma_F.C.
- 20: 520651 U.S._Città_di_Palermo

21: 2358 A.S._Roma
22: 18951905 Argentina
23: 1687460 Carlos_Bianchi
24: 14532 Italy
25: 3319700 Arrigo_Sacchi
26: 1138777 A.C._Cesena
27: 25458 Rome
28: 149877 Udinese_Calcio
29: 11867 Germany
30: 1129681 Oliver_Bierhoff
31: 15496 Serie_A
32: 1691772 U.S._Cremonese
33: 18940588 A.C._Milan
34: 31843 Uruguay
35: 9158864 Óscar_Tabárez
36: 3165 ACF_Fiorentina
37: 44220 UEFA_Champions_League
38: 16415 Juventus_F.C.
39: 3073616 A.S.D._Fidelis_Andria_1928
40: 15496 Serie_A
41: 1136763 S.S.D._Piacenza_Calcio_1919
42: 210482 Atalanta_B.C.
43: 1135354 U.S._Lecce
44: 30872779 Genoa_C.F.C.
45: 30872779 Genoa_C.F.C.
46: 574307 U.C._Sampdoria
47: 3085203 A.S.G._Nocerina
48: 1136763 S.S.D._Piacenza_Calcio_1919
49: 15496 Serie_A
50: 2592452 A.C._Perugia_Calcio
1: 4918282 1996-97_UEFA_Cup
 [winner: 232175 UEFA_Europa_League, pem_pos=0,
final_scores=0.33266228, final_local_scores=0.31710753,
pem=-0.260892, similarity=14.971605]
 gt: 4918282 1996-97_UEFA_Cup, pem_pos=15,
final_scores=-0.18086885, final_local_scores=-0.15856636,
pem=-3.572804, similarity=18.787241]
2: 1281200 Coppa_Italia
 [winner: 11237 FA_Cup, pem_pos=6,
final_scores=-0.15056506, final_local_scores=-0.16035241,
pem=-3.1209865, similarity=16.276886]
 RECALL MISS: 1281200 Coppa_Italia]
3: 4462898 History_of_Cosenza_Calcio_1914
 [winner: 36183605 Cosenza_Calcio, pem_pos=1,
final_scores=-0.15667686, final_local_scores=-0.23381928,
pem=-0.63829803, similarity=0.58324313]
 RECALL MISS: 4462898 History_of_Cosenza_Calcio_1914]
1: 62611 Parma_F.C.
 [pem_pos=1, final_scores=0.6508159,
final_local_scores=0.6062872, pem=-1.0853336,
similarity=28.084743]
2: 2358 A.S._Roma

```
[pem_pos=0, final_scores=0.7235161,
final_local_scores=0.68358123, pem=-1.2216192,
similarity=31.132498]
3: 149877 Udinese_Calcio
    [pem_pos=0, final_scores=0.7803505,
final_local_scores=0.7571472, pem=0.0, similarity=26.684496]
4: 1281200 Coppa_Italia
    [pem_pos=0, final_scores=0.44373804,
final_local_scores=0.40173376, pem=0.0, similarity=16.078249]
5: 25458 Rome
    [pem_pos=0, final_scores=0.49606025,
final_local_scores=0.5539021, pem=-0.2517334,
similarity=21.98822]
6: 62611 Parma_F.C.
    [pem_pos=1, final_scores=0.81543785,
final_local_scores=0.7861117, pem=-1.0853336,
similarity=33.451065]
7: 2358 A.S._Roma
    [pem_pos=0, final_scores=0.54707384,
final_local_scores=0.49084485, pem=-1.2216192,
similarity=25.380857]
8: 1281200 Coppa_Italia
    [pem_pos=0, final_scores=0.36791134,
final_local_scores=0.3189046, pem=0.0, similarity=13.6064625]
9: 18940588 A.C._Milan
    [pem_pos=1, final_scores=0.78855056,
final_local_scores=0.7598714, pem=-1.9620993,
similarity=37.435993]
10: 1136866 Empoli_F.C.
    [pem_pos=0, final_scores=0.50681967,
final_local_scores=0.47060364, pem=0.0, similarity=18.133465]
11: 62611 Parma_F.C.
    [pem_pos=1, final_scores=0.8072055,
final_local_scores=0.77711904, pem=-1.0853336,
similarity=33.18271]
12: 14532 Italy
    [pem_pos=0, final_scores=0.629016,
final_local_scores=0.68808043, pem=-0.75948656,
similarity=28.75361]
13: 1142794 Carlo_Ancelotti
    [pem_pos=0, final_scores=0.84421355,
final_local_scores=0.88034815, pem=0.0, similarity=30.361057]
14: 1527955 Enrico_Chiesa
    [pem_pos=0, final_scores=0.74398917,
final_local_scores=0.77808654, pem=0.0, similarity=27.309368]
15: 62611 Parma_F.C.
    [pem_pos=1, final_scores=0.6821806,
final_local_scores=0.64054835, pem=-1.0853336,
similarity=29.107166]
16: 62611 Parma_F.C.
    [pem_pos=1, final_scores=0.7315269,
final_local_scores=0.6944517, pem=-1.0853336,
```

```
similarity=30.71575]
17: 18424280 Alessandro_Melli
    [pem_pos=0, final_scores=0.5651956,
final_local_scores=0.59741616, pem=0.0, similarity=21.917803]
18: 15496 Serie_A
    [pem_pos=0, final_scores=0.4576162,
final_local_scores=0.38810492, pem=-0.1387667,
similarity=16.426172]
19: 62611 Parma_F.C.
    [pem_pos=1, final_scores=0.8051465,
final_local_scores=0.7748699, pem=-1.0853336,
similarity=33.11559]
20: 520651 U.S._Città_di_Palermo
    [pem_pos=1, final_scores=0.6219482,
final_local_scores=0.5766, pem=-1.449244,
similarity=29.177822]
21: 2358 A.S._Roma
    [pem_pos=0, final_scores=0.62366337,
final_local_scores=0.5745073, pem=-1.2216192,
similarity=27.877512]
22: 18951905 Argentina
    [pem_pos=0, final_scores=0.5932392,
final_local_scores=0.68054163, pem=-0.18885723,
similarity=25.425468]
23: 1687460 Carlos_Bianchi
    [pem_pos=0, final_scores=0.5555517,
final_local_scores=0.58430487, pem=0.0, similarity=21.526535]
24: 14532 Italy
    [pem_pos=0, final_scores=0.5196004,
final_local_scores=0.56856054, pem=-0.75948656,
similarity=25.1869]
25: 3319700 Arrigo_Sacchi
    [pem_pos=0, final_scores=0.7878757,
final_local_scores=0.8253227, pem=0.0, similarity=28.718988]
26: 1138777 A.C._Cesena
    [pem_pos=1, final_scores=0.5238571,
final_local_scores=0.49324304, pem=-0.054029576,
similarity=19.10289]
27: 25458 Rome
    [pem_pos=0, final_scores=0.44811708,
final_local_scores=0.5015315, pem=-0.2517334,
similarity=20.425377]
28: 149877 Udinese_Calcio
    [pem_pos=0, final_scores=0.5329233,
final_local_scores=0.4868706, pem=0.0, similarity=18.618902]
29: 11867 Germany
    [pem_pos=0, final_scores=0.6380756,
final_local_scores=0.7555702, pem=-0.15529114,
similarity=27.481932]
30: 1129681 Oliver_Bierhoff
    [pem_pos=0, final_scores=0.9608601,
final_local_scores=1.0361567, pem=0.0, similarity=35.010696]
```

31: 15496 Serie_A
[pem_pos=0, final_scores=0.6217886,
final_local_scores=0.5674383, pem=-0.1387667,
similarity=21.77784]
32: 1691772 U.S. Cremonese
[pem_pos=0, final_scores=0.36310428,
final_local_scores=0.32387814, pem=0.0, similarity=13.754883]
33: 18940588 A.C. Milan
[pem_pos=1, final_scores=0.6875568,
final_local_scores=0.6495511, pem=-1.9620993,
similarity=34.143814]
34: 31843 Uruguay
[pem_pos=0, final_scores=0.50337654,
final_local_scores=0.57574624, pem=-0.065879524,
similarity=21.629393]
35: 9158864 Óscar Tabárez
[pem_pos=0, final_scores=0.6167793,
final_local_scores=0.674775, pem=0.0, similarity=24.226345]
36: 3165 ACF Fiorentina
[pem_pos=0, final_scores=0.49847466,
final_local_scores=0.43850368, pem=0.0, similarity=17.175537]
37: 44220 UEFA Champions League
[pem_pos=0, final_scores=0.3615486,
final_local_scores=0.33813307, pem=-0.5339415,
similarity=17.083935]
38: 16415 Juventus F.C.
[pem_pos=0, final_scores=0.51678497,
final_local_scores=0.4607867, pem=0.0, similarity=17.840508]
39: 3073616 A.S.D. Fidelis Andria 1928
[pem_pos=0, final_scores=0.5261517,
final_local_scores=0.54311913, pem=0.0, similarity=20.297472]
40: 15496 Serie_A
[pem_pos=0, final_scores=0.55928475,
final_local_scores=0.49916238, pem=-0.1387667,
similarity=19.740349]
41: 1136763 S.S.D. Piacenza Calcio 1919
[pem_pos=1, final_scores=0.5091737,
final_local_scores=0.47323853, pem=-1.1110737,
similarity=24.254288]
42: 210482 Atalanta B.C.
[pem_pos=0, final_scores=0.5825791,
final_local_scores=0.53767294, pem=-0.3467802,
similarity=22.02079]
43: 1135354 U.S. Lecce
[pem_pos=1, final_scores=0.5055116,
final_local_scores=0.45890802, pem=-0.4613229,
similarity=20.29319]
44: 30872779 Genoa C.F.C.
[pem_pos=1, final_scores=0.6923222,
final_local_scores=0.6525307, pem=-1.6777803,
similarity=32.68656]
45: 30872779 Genoa C.F.C.

```
[pem_pos=1, final_scores=0.74736816,
final_local_scores=0.71266013, pem=-1.6777803,
similarity=34.48094]
46: 574307 U.C._Sampdoria
    [pem_pos=0, final_scores=0.50647956,
final_local_scores=0.45070416, pem=0.0, similarity=17.539623]
47: 3085203 A.S.G._Nocerina
    [pem_pos=0, final_scores=0.4881587,
final_local_scores=0.51487917, pem=0.0, similarity=19.454735]
48: 1136763 S.S.D._Piacenza_Calcio_1919
    [pem_pos=1, final_scores=0.4817013,
final_local_scores=0.44322914, pem=-1.1110737,
similarity=23.358746]
49: 15496 Serie_A
    [pem_pos=0, final_scores=0.5644495,
final_local_scores=0.504804, pem=-0.1387667,
similarity=19.908707]
50: 2592452 A.C._Perugia_Calcio
    [pem_pos=1, final_scores=0.6027207,
final_local_scores=0.56551564, pem=-0.7535512,
similarity=25.063757]
1: 232175 UEFA_Europa_League
    [pem_pos=0, final_scores=0.33266228,
final_local_scores=0.31710753, pem=-0.260892,
similarity=14.971605]
2: 1411049 Delfino_Pescara_1936
    [pem_pos=1, final_scores=0.05684711,
final_local_scores=-0.01940246, pem=-0.7725453,
similarity=7.7119265]
```

Sample: b'674_SOCCER'

SOCCER - [1 1 BAYERN 1] HIT FOUR TO TAKE [2 2 BUNDESLIGA 2] TOP SPOT . [3 3 BONN 3] 1996-08-28 Goals from [4 4 Thomas Helmer 4] and [5 5 Juergen Klinsmann 5] helped [6 6 Bayern Munich 6] to a 4-2 home win over [7 7 Bayer Leverkusen 7] on Wednesday and powered them to the top of the [8 8 Bundesliga 8] . The comfortable victory gave [9 9 Bayern 9] 10 points from their first four games , a point ahead of second-placed [10 10 Stuttgart 10] , who have a game in hand . [11 11 Brazilian 11] midfielder [12 12 Paulo Sergio 12] put [13 13 Leverkusen 13] ahead in the 25th minute but [14 14 Alexander Zickler 14] equalised just a minute later . A header from [15 15 Helmer 15] and an acrobatic strike from [16 16 Klinsmann 16] gave [17 17 Bayern 17] a two-goal cushion at halftime . But the pick of the 13-times champions ' goals came from [18 18 Ruggiero Rizzitelli 18] , who beat three defenders to put [19 19 Bayern 19] 4-1 up . [20 20 Markus Feldhoff 20] hit a consolation goal for [21 21 Leverkusen 21] . [22 22 Hansa Rostock 22] brought [23 23 Cologne 23] 's 100 percent record to an end with a 2-0 win over the Rhineside club while a [24 24 Sean Dundee 24] hat-trick inside seven minutes stood out in [25 25 Karlsruhe 25] 's 4-0 demolition of [26 26 St Pauli 26] .

```
1: 172326 FC_Bayern_Munich
2: 686142 Bundesliga
3: 3295 Bonn
4: 2568262 Thomas_Helmer
5: 525141 Jürgen_Klinsmann
6: 172326 FC_Bayern_Munich
7: 502005 Bayer_04_Leverkusen
8: 686142 Bundesliga
9: 172326 FC_Bayern_Munich
10: 737608 VfB_Stuttgart
11: 3383 Brazil
12: 5071886 Paulo_Sérgio_Silvestre_do_Nascimento
13: 502005 Bayer_04_Leverkusen
14: 2431610 Alexander_Zickler
15: 2568262 Thomas_Helmer
16: 525141 Jürgen_Klinsmann
17: 172326 FC_Bayern_Munich
18: 7560958 Ruggiero_Rizzitelli
19: 172326 FC_Bayern_Munich
20: 14466734 Markus_Feldhoff
21: 502005 Bayer_04_Leverkusen
22: 599978 F.C._Hansa_Rostock
23: 497888 1._FC_Köln
24: 3373736 Sean_Dundee
25: 1734592 Karlsruher_SC
26: 881244 FC_St._Pauli
1: 172326 FC_Bayern_Munich
    [pem_pos=1, final_scores=0.5557155,
final_local_scores=0.4998837, pem=-1.5624554,
similarity=27.504114]
2: 686142 Bundesliga
    [pem_pos=0, final_scores=0.4655376,
final_local_scores=0.41306764, pem=-0.08715299,
similarity=16.890427]
3: 3295 Bonn
    [pem_pos=0, final_scores=0.35434645,
final_local_scores=0.4343832, pem=0.0, similarity=17.052574]
4: 2568262 Thomas_Helmer
    [pem_pos=0, final_scores=0.9580651,
final_local_scores=1.0133727, pem=0.0, similarity=34.330776]
5: 525141 Jürgen_Klinsmann
    [pem_pos=0, final_scores=0.86524826,
final_local_scores=0.8744735, pem=0.0, similarity=30.185745]
6: 172326 FC_Bayern_Munich
    [pem_pos=0, final_scores=0.4851666,
final_local_scores=0.4228198, pem=0.0, similarity=16.707499]
7: 502005 Bayer_04_Leverkusen
    [pem_pos=0, final_scores=0.5999947,
final_local_scores=0.5501154, pem=0.0, similarity=20.506254]
8: 686142 Bundesliga
    [pem_pos=0, final_scores=0.3752169,
final_local_scores=0.31440604, pem=-0.08715299,
```

```
similarity=13.946168]
9: 172326 FC_Bayern_Munich
    [pem_pos=1, final_scores=0.49279195,
final_local_scores=0.4311493, pem=-1.5624554,
similarity=25.452944]
10: 737608 VfB_Stuttgart
    [pem_pos=1, final_scores=0.59905034,
final_local_scores=0.5578473, pem=-2.1952448,
similarity=32.67507]
11: 3383 Brazil
    [pem_pos=0, final_scores=0.46829098,
final_local_scores=0.58656746, pem=-0.2733083,
similarity=23.080349]
12: 5071886 Paulo_Sérgio_Silvestre_do_Nascimento
    [pem_pos=0, final_scores=0.45089585,
final_local_scores=0.5139013, pem=0.0, similarity=19.425552]
13: 502005 Bayer_04_Leverkusen
    [pem_pos=1, final_scores=0.43503195,
final_local_scores=0.36991864, pem=-1.2305675,
similarity=21.82084]
14: 2431610 Alexander_Zickler
    [pem_pos=0, final_scores=0.8959463,
final_local_scores=0.93232435, pem=0.0, similarity=31.912132]
15: 2568262 Thomas_Helmer
    [pem_pos=0, final_scores=0.90928346,
final_local_scores=0.9682209, pem=0.0, similarity=32.983356]
16: 525141 Jürgen_Klinsmann
    [pem_pos=0, final_scores=0.59314054,
final_local_scores=0.5772371, pem=0.0, similarity=21.31562]
17: 172326 FC_Bayern_Munich
    [pem_pos=1, final_scores=0.48701936,
final_local_scores=0.4248436, pem=-1.5624554,
similarity=25.264769]
18: 7560958 Ruggiero_Rizzitelli
    [pem_pos=0, final_scores=0.4084491,
final_local_scores=0.4968434, pem=0.0, similarity=18.916512]
19: 172326 FC_Bayern_Munich
    [pem_pos=1, final_scores=0.44175977,
final_local_scores=0.37540442, pem=-1.5624554,
similarity=23.789406]
20: 14466734 Markus_Feldhoff
    [pem_pos=0, final_scores=0.7304909,
final_local_scores=0.8540787, pem=0.0, similarity=29.577126]
21: 502005 Bayer_04_Leverkusen
    [pem_pos=1, final_scores=0.4573459,
final_local_scores=0.39429325, pem=-1.2305675,
similarity=22.548227]
22: 599978 F.C._Hansa_Rostock
    [pem_pos=0, final_scores=0.54577905,
final_local_scores=0.5206973, pem=0.0, similarity=19.628359]
23: 497888 1._FC_Köln
    [pem_pos=10, final_scores=0.8436598,
```

```
final_local_scores=0.8245008, pem=-3.09805,
similarity=45.54214]
24: 3373736 Sean_Dundee
    [pem_pos=0, final_scores=0.47343302,
final_local_scores=0.52894765, pem=0.0, similarity=19.874565]
25: 1734592 Karlsruher_SC
    [pem_pos=1, final_scores=0.6176569,
final_local_scores=0.58323175, pem=-2.2263577,
similarity=33.60179]
26: 881244 FC_St._Pauli
    [pem_pos=1, final_scores=0.5785592,
final_local_scores=0.5609268, pem=-0.11686432,
similarity=21.464415]
```

Sample: b'675_SOCCER'

```
SOCCER - [1 1 ITALIAN CUP 1] SECOND ROUND RESULTS . [2 2 ROME
2] 1996-08-28 Results of [3 3 Italian Cup 3] second round
matches played on Wednesday : [4 4 Empoli 4] 1 [5 5 Milan 5]
1 Spal 2 [6 6 Reggiana 6] 4 [7 7 Lucchese 7] 1 [8 8 Vicenza 8]
2 [9 9 Cremonese 9] 2 [10 10 Udinese 10] 1 [11 11 Cesena 11]
3 [12 12 Roma 12] 1 [13 13 Bologna 13] 2 [14 14 Torino 14] 1 1
Cosenza 1 [15 15 Fiorentina 15] 3 Avellino 0 [16 16 Lazio 16]
1 [17 17 Bari 17] 1 [18 18 Verona 18] 1 [19 19 Pescara 19] 3 [
20 20 Parma 20] 1 [21 21 Monza 21] 0 [22 22 Napoli 22] 1 [23 2
3 Chievo 23] 2 [24 24 Cagliari 24] 3 [25 25 Ravenna 25] 0 [26
26 Inter 26] 1 [27 27 Fidelis Andria 27] 0 [28 28 Juventus 28]
2
```

```
1: 1281200 Coppa_Italia
2: 25458 Rome
3: 1281200 Coppa_Italia
4: 1136866 Empoli_F.C.
5: 18940588 A.C._Milan
6: 3070366 A.C._Reggiana_1919
7: 33707714 A.S._Lucchese_Libertas_1905
8: 1136820 Vicenza_Calcio
9: 1691772 U.S._Cremonese
10: 149877 Udinese_Calcio
11: 1138777 A.C._Cesena
12: 2358 A.S._Roma
13: 81238 Bologna_F.C._1909
14: 19450750 Torino_F.C.
15: 3165 ACF_Fiorentina
16: 28984 S.S._Lazio
17: 1138693 A.S._Bari
18: 14035 Hellas_Verona_F.C.
19: 1411049 Delfino_Pescara_1936
20: 62611 Parma_F.C.
21: 2870603 A.C._Monza_Brianza_1912
22: 1130620 S.S.C._Napoli
23: 6903 A.C._Chievo_Verona
24: 1116811 Cagliari_Calcio
25: 2876568 Ravenna_F.C.
```

```
26: 15116 Inter_Milan
27: 3073616 A.S.D._Fidelis_Andria_1928
28: 16415 Juventus_F.C.
1: 4462898 History_of_Cosenza_Calcio_1914
    [winner: 36183605 Cosenza_Calcio, pem_pos=1,
final_scores=-0.16000688, final_local_scores=-0.24937348,
pem=-0.63829803, similarity=0.11907476]
    RECALL MISS: 4462898 History_of_Cosenza_Calcio_1914]
1: 1281200 Coppa_Italia
    [pem_pos=0, final_scores=0.43453807,
final_local_scores=0.3930443, pem=0.0, similarity=15.818938]
2: 25458 Rome
    [pem_pos=0, final_scores=0.4309404,
final_local_scores=0.50889665, pem=-0.2517334,
similarity=20.645168]
3: 1281200 Coppa_Italia
    [pem_pos=0, final_scores=0.36348194,
final_local_scores=0.3154263, pem=0.0, similarity=13.502663]
4: 1136866 Empoli_F.C.
    [pem_pos=0, final_scores=0.5212311,
final_local_scores=0.47676772, pem=0.0, similarity=18.317413]
5: 18940588 A.C._Milan
    [pem_pos=1, final_scores=0.58210593,
final_local_scores=0.54161745, pem=-1.9620993,
similarity=30.922861]
6: 3070366 A.C._Reggiana_1919
    [pem_pos=0, final_scores=0.5156542,
final_local_scores=0.48006976, pem=0.0, similarity=18.415953]
7: 33707714 A.S._Lucchese_Libertas_1905
    [pem_pos=0, final_scores=0.56829244,
final_local_scores=0.54373497, pem=-0.15317558,
similarity=21.148842]
8: 1136820 Vicenza_Calcio
    [pem_pos=1, final_scores=0.4953463,
final_local_scores=0.44782406, pem=-0.7035717,
similarity=21.279808]
9: 1691772 U.S._Cremonese
    [pem_pos=0, final_scores=0.5360289,
final_local_scores=0.503958, pem=0.0, similarity=19.128824]
10: 149877 Udinese_Calcio
    [pem_pos=0, final_scores=0.61858505,
final_local_scores=0.5825071, pem=0.0, similarity=21.472885]
11: 1138777 A.C._Cesena
    [pem_pos=1, final_scores=0.47358924,
final_local_scores=0.42852414, pem=-0.054029576,
similarity=17.171549]
12: 2358 A.S._Roma
    [pem_pos=0, final_scores=0.52981716,
final_local_scores=0.4735554, pem=-1.2216192,
similarity=24.864906]
13: 81238 Bologna_F.C._1909
    [pem_pos=1, final_scores=0.5635833,
```

```
final_local_scores=0.51195544, pem=-1.6631861,
similarity=28.41215]
14: 19450750 Torino_F.C.
    [pem_pos=0, final_scores=0.60949343,
final_local_scores=0.56606174, pem=-0.52662677,
similarity=23.846003]
15: 3165 ACF_Fiorentina
    [pem_pos=0, final_scores=0.52826625,
final_local_scores=0.4716075, pem=0.0, similarity=18.163422]
16: 28984 S.S._Lazio
    [pem_pos=1, final_scores=0.44952422,
final_local_scores=0.39133513, pem=-0.6322365,
similarity=19.206135]
17: 1138693 A.S._Bari
    [pem_pos=1, final_scores=0.81778365,
final_local_scores=0.79097134, pem=-1.3113433,
similarity=34.825165]
18: 14035 Hellas_Verona_F.C.
    [pem_pos=1, final_scores=0.8028386,
final_local_scores=0.7748317, pem=-2.0420845,
similarity=38.31741]
19: 1411049 Delfino_Pescara_1936
    [pem_pos=1, final_scores=0.41310954,
final_local_scores=0.35866708, pem=-0.7725453,
similarity=18.994274]
20: 62611 Parma_F.C.
    [pem_pos=1, final_scores=0.83664626,
final_local_scores=0.8089181, pem=-1.0853336,
similarity=34.131657]
21: 2870603 A.C._Monza_Brianza_1912
    [pem_pos=2, final_scores=0.4402961,
final_local_scores=0.40690273, pem=-1.8445003,
similarity=26.26318]
22: 1130620 S.S.C._Napoli
    [pem_pos=0, final_scores=0.59560716,
final_local_scores=0.54884106, pem=-0.37227353,
similarity=22.492706]
23: 6903 A.C._Chievo_Verona
    [pem_pos=0, final_scores=0.5424469,
final_local_scores=0.49369222, pem=0.0, similarity=18.822474]
24: 1116811 Cagliari_Calcio
    [pem_pos=1, final_scores=0.4028718,
final_local_scores=0.34565723, pem=-0.8688384,
similarity=19.129692]
25: 2876568 Ravenna_F.C.
    [pem_pos=1, final_scores=0.47308463,
final_local_scores=0.45107192, pem=-1.9313984,
similarity=28.053844]
26: 15116 Inter_Milan
    [pem_pos=0, final_scores=0.4266371,
final_local_scores=0.36525732, pem=-0.015552434,
similarity=15.074297]
```

```
27: 3073616 A.S.D._Fidelis_Andria_1928
    [pem_pos=0, final_scores=0.54627365,
final_local_scores=0.55261177, pem=0.0, similarity=20.58075]
28: 16415 Juventus_F.C.
    [pem_pos=0, final_scores=0.43617398,
final_local_scores=0.37948585, pem=0.0, similarity=15.414328]
```

Sample: b'676_SOCCER'

SOCCER - [1 1 GERMAN 1] FIRST DIVISION SUMMARIES . [2 2 BONN

2] 1996-08-28 Summaries of Wednesday 's [3 3 German 3] first division soccer matches : [4 4 Karlsruhe 4] 4 (1 Keller 18th minute , 2 Dundee 56th 59th and 64th) [5 5 St Pauli 5] 0 . Halftime 1-0 . Attendance 27,600 . [6 6 Bayern Munich 6] 4 ([7 7 Zickler 7] 26th , [8 8 Helmer 8] 37th , [9 9 Klinsmann 9] 44th , [10 10 Rizzitelli 10] 48th) [11 11 Bayer Leverkusen 11] 2 (Sergio 25th , [12 12 Feldhoff 12] 54th) . 3-1 . 48,000 . [13 13 Cologne 13] 0 [14 14 Hansa Rostock 14] 2 ([15 15 Akpoborie 15] 5th and 59th) . 0-1 . 27,000 . [16 16 Fortuna Duesseldorf 16] 0 [17 17 1860 Munich 17] 0 . 11,500 . [18 18 Arminia Bielefeld 18] 1 (Von Heesen 56th) [19 19 Duisburg 19] 1 (3 Hirsch 65th) . 0-0 . 15,000 .

1: 11867 Germany

2: 3295 Bonn

3: 11867 Germany

4: 1734592 Karlsruher SC

5: 881244 FC_St._Pauli

6: 172326 FC_Bayern_Munich

7: 2431610 Alexander_Zickler

8: 2568262 Thomas_Helmer

9: 525141 Jürgen_Klinsmann

10: 7560958 Ruggiero_Rizzitelli

11: 502005 Bayer_04_Leverkusen

12: 14466734 Markus_Feldhoff

13: 497888 1._FC_Köln

14: 599978 F.C._Hansa_Rostock

15: 9649414 Jonathan_Akpoborie

16: 1949688 Fortuna_Düsseldorf

17: 1007441 TSV_1860_München

18: 1145228 Arminia_Bielefeld

19: 1586579 MSV_Duisburg

1: 10634409 Marc_Keller

[winner: 895763 Kasey_Keller, pem_pos=3,
final_scores=-0.062668875, final_local_scores=-0.0584116,
pem=-2.8574977, similarity=17.886112 |
RECALL MISS: 10634409 Marc_Keller]

2: 3373736 Sean_Dundee

[winner: 306504 Dundee_F.C., pem_pos=1,
final_scores=-0.06081797, final_local_scores=-0.07314503,
pem=-1.3661808, similarity=9.336423 |
RECALL MISS: 3373736 Sean_Dundee]

3: 21064133 Dietmar_Hirsch

[winner: 10924639 Julius_Hirsch, pem_pos=5,

```
final_scores=-0.270805, final_local_scores=-0.24844995,
pem=-3.2767081, similarity=14.494719]
    RECALL MISS: 21064133 Dietmar_Hirsch]
1: 11867 Germany
    [pem_pos=1, final_scores=0.79202646,
final_local_scores=0.8625755, pem=-1.0465823,
similarity=35.522163]
2: 3295 Bonn
    [pem_pos=0, final_scores=0.38195586,
final_local_scores=0.44212496, pem=0.0, similarity=17.283604]
3: 11867 Germany
    [pem_pos=1, final_scores=0.56609076,
final_local_scores=0.61577505, pem=-1.0465823,
similarity=28.157146]
4: 1734592 Karlsruher_SC
    [pem_pos=1, final_scores=0.7713453,
final_local_scores=0.748771, pem=-2.2263577,
similarity=38.541813]
5: 881244 FC_St._Pauli
    [pem_pos=1, final_scores=0.56590825,
final_local_scores=0.53850883, pem=-0.11686432,
similarity=20.795418]
6: 172326 FC_Bayern_Munich
    [pem_pos=0, final_scores=0.47243518,
final_local_scores=0.42292958, pem=0.0, similarity=16.710775]
7: 2431610 Alexander_Zickler
    [pem_pos=0, final_scores=0.6545606,
final_local_scores=0.67184955, pem=0.0, similarity=24.139044]
8: 2568262 Thomas_Helmer
    [pem_pos=2, final_scores=0.5636106,
final_local_scores=0.60546774, pem=-1.4827579,
similarity=30.221546]
9: 525141 Jürgen_Klinsmann
    [pem_pos=0, final_scores=0.509823,
final_local_scores=0.5036977, pem=0.0, similarity=19.121058]
10: 7560958 Ruggiero_Rizzitelli
    [pem_pos=0, final_scores=0.43161184,
final_local_scores=0.5310946, pem=0.0, similarity=19.938635]
11: 502005 Bayer_04_Leverkusen
    [pem_pos=0, final_scores=0.44663006,
final_local_scores=0.39149857, pem=0.0, similarity=15.772811]
12: 14466734 Markus_Feldhoff
    [pem_pos=0, final_scores=0.5865455,
final_local_scores=0.6980339, pem=-0.6931472,
similarity=28.689878]
13: 497888 1._FC_Köln
    [pem_pos=10, final_scores=0.66621864,
final_local_scores=0.63145053, pem=-3.09805,
similarity=39.781136]
14: 599978 F.C._Hansa_Rostock
    [pem_pos=0, final_scores=0.58331424,
final_local_scores=0.5529238, pem=0.0, similarity=20.590061]
```

```
15: 9649414 Jonathan_Akpoborie
    [pem_pos=0, final_scores=0.61174756,
final_local_scores=0.68823814, pem=0.0, similarity=24.628113]
16: 1949688 Fortuna_Düsseldorf
    [pem_pos=0, final_scores=0.5807175,
final_local_scores=0.54612243, pem=0.0, similarity=20.387096]
17: 1007441 TSV_1860_München
    [pem_pos=0, final_scores=0.44980973,
final_local_scores=0.40285593, pem=0.0, similarity=16.111738]
18: 1145228 Arminia_Bielefeld
    [pem_pos=0, final_scores=0.46942002,
final_local_scores=0.42757368, pem=0.0, similarity=16.849365]
19: 1586579 MSV_Duisburg
    [pem_pos=1, final_scores=0.4889204,
final_local_scores=0.45455867, pem=-1.5900439,
similarity=26.301556]
```

Sample: b'677_SOCCER'

SOCCER - LEADING [1 1 FRENCH 1] LEAGUE SCORERS . [2 2 PARIS 2] 1996-08-28 Leading scorers in the [3 3 French 3] first division after Wednesday 's matches : 3 - Anton Drobniak ([4 4 Bastia 4]) , [5 5 Vladimir Smicer 5] ([6 6 Lens 6]) , [7 7 Miladin Bećanović 7] ([8 8 Lille 8]) , [9 9 Alain Caveglia 9] ([10 10 Lyon 10]) , [11 11 Xavier Gravelaine 11] ([12 12 Marseille 12]) , [13 13 Robert Pires 13] ([14 14 Metz 14]) , [15 15 Thierry Henry 15] ([16 16 Monaco 16]) 2 - [17 17 Christopher Wreh 17] ([18 18 Guingamp 18]) , [19 19 Marc-Vivien Foe 19] ([20 20 Lens 20]) , [21 21 Enzo Scifo 21] ([22 22 Monaco 22]) , [23 23 James Debbah 23] ([24 24 Nice 24]) , [25 25 Patrice Loko 25] ([26 26 PSG 26]) , [27 27 Stéphane Guivarch 27] ([28 28 Rennes 28])

```
1: 5843419 France
2: 22989 Paris
3: 5843419 France
4: 1256442 SC_Bastia
5: 549482 Vladimír Šmicer
6: 1075697 RC_Lens
7: 13449225 Miladin_Bećanović
8: 990757 Lille_OSC
9: 11576169 Alain_Caveglia
10: 979521 Olympique_Lyonnais
11: 10326208 Xavier_Gravelaine
12: 492068 Olympique_de_Marseille
13: 553772 Robert_Pirès
14: 1135310 FC_Metz
15: 220160 Thierry_Henry
16: 222981 AS_Monaco_FC
17: 2839275 Christopher_Wreh
18: 1693070 En_Avant_de_Guingamp
19: 253764 Marc-Vivien_Foé
20: 1075697 RC_Lens
21: 52265 Enzo_Scifo
```

```
22: 222981 AS_Monaco_FC
23: 11527631 James_Débbah
24: 704478 OGC_Nice
25: 3254011 Patrice_Loko
26: 357488 Paris_Saint-Germain_F.C.
27: 1435187 Stéphane_Guivarc'h
28: 1004187 Stade_Rennais_F.C.
1: 5843419 France
    [pem_pos=0, final_scores=0.6285725,
final_local_scores=0.6961029, pem=-0.80223274,
similarity=29.22548]
2: 22989 Paris
    [pem_pos=0, final_scores=0.5103446,
final_local_scores=0.57327473, pem=-0.115301244,
similarity=21.8244]
3: 5843419 France
    [pem_pos=0, final_scores=0.6091159,
final_local_scores=0.6748495, pem=-0.80223274,
similarity=28.591234]
4: 1256442 SC_Bastia
    [pem_pos=1, final_scores=0.48293525,
final_local_scores=0.44329005, pem=-0.7022244,
similarity=21.137177]
5: 549482 Vladimír_Šmicer
    [pem_pos=0, final_scores=0.8840062,
final_local_scores=0.981993, pem=0.0, similarity=33.394344]
6: 1075697 RC_Lens
    [pem_pos=0, final_scores=0.5574445,
final_local_scores=0.51399904, pem=-0.86081105,
similarity=24.109695]
7: 13449225 Miladin_Bečanović
    [pem_pos=0, final_scores=0.3756293,
final_local_scores=0.5780808, pem=0.0, similarity=21.340796]
8: 990757 Lille OSC
    [pem_pos=1, final_scores=0.6002643,
final_local_scores=0.5609146, pem=-1.3501564,
similarity=28.170883]
9: 11576169 Alain_Caveglia
    [pem_pos=0, final_scores=0.6852228,
final_local_scores=0.77576524, pem=0.0, similarity=27.240095]
10: 979521 Olympique_Lyonnais
    [pem_pos=1, final_scores=0.5444377,
final_local_scores=0.49391323, pem=-1.9193003,
similarity=29.266521]
11: 10326208 Xavier_Gravelaine
    [pem_pos=0, final_scores=0.58822364,
final_local_scores=0.6738104, pem=0.0, similarity=24.19756]
12: 492068 Olympique_de_Marseille
    [pem_pos=1, final_scores=0.595614,
final_local_scores=0.54574865, pem=-1.4345005,
similarity=28.176977]
13: 553772 Robert_Pirès
```

[pem_pos=0, final_scores=0.8987251,
final_local_scores=0.9537901, pem=0.0, similarity=32.55271]
14: 1135310 FC_Metz
[pem_pos=1, final_scores=0.62342757,
final_local_scores=0.5962805, pem=-1.9220247,
similarity=32.33618]
15: 220160 Thierry_Henry
[pem_pos=0, final_scores=0.73722297,
final_local_scores=0.7647676, pem=0.0, similarity=26.911903]
16: 222981 AS_Monaco_FC
[pem_pos=1, final_scores=0.82434696,
final_local_scores=0.80306596, pem=-2.451784,
similarity=41.387985]
17: 2839275 Christopher_Wreh
[pem_pos=0, final_scores=0.6864576,
final_local_scores=0.7522196, pem=0.0, similarity=26.537447]
18: 1693070 En_Avant_de_Guingamp
[pem_pos=0, final_scores=0.46577746,
final_local_scores=0.4380859, pem=0.0, similarity=17.16307]
19: 253764 Marc-Vivien_Foé
[pem_pos=0, final_scores=0.7612185,
final_local_scores=0.82396495, pem=0.0, similarity=28.67847]
20: 1075697 RC_Lens
[pem_pos=0, final_scores=0.48642057,
final_local_scores=0.4364162, pem=-0.86081105,
similarity=21.794468]
21: 52265 Enzo_Scifo
[pem_pos=0, final_scores=1.0078827,
final_local_scores=1.080966, pem=0.0, similarity=36.347893]
22: 222981 AS_Monaco_FC
[pem_pos=1, final_scores=0.70502335,
final_local_scores=0.6727231, pem=-2.451784,
similarity=37.498295]
23: 11527631 James_Debbah
[pem_pos=0, final_scores=0.55815935,
final_local_scores=0.6369928, pem=0.0, similarity=23.098848]
24: 704478 OGC_Nice
[pem_pos=1, final_scores=0.61229354,
final_local_scores=0.5831274, pem=-2.1390686,
similarity=33.123985]
25: 3254011 Patrice_Loko
[pem_pos=0, final_scores=0.6479637,
final_local_scores=0.769595, pem=0.0, similarity=27.055964]
26: 357488 Paris_Saint-Germain_F.C.
[pem_pos=0, final_scores=0.5190178,
final_local_scores=0.46684223, pem=-0.4638402,
similarity=20.543652]
27: 1435187 Stéphane_Guivarc'h
[pem_pos=0, final_scores=0.8083077,
final_local_scores=0.8447702, pem=0.0, similarity=29.299341]
28: 1004187 Stade_Rennais_F.C.
[pem_pos=1, final_scores=0.42103738,

```
final_local_scores=0.3659395, pem=-1.0138907,  
similarity=20.523773]
```

Sample: b'678_SOCCER'

SOCCER - [1 1 SMICER 1] 'S LAST-GASP GOAL KEEPS [2 2 LENS 2] IN THE LEAD . [3 3 PARIS 3] 1996-08-28 [4 4 Euro 96 4] star [5 5 Vladimir Smicer 5] of the [6 6 Czech Republic 6] scored at the last second for [7 7 Lens 7] , allowing them to retain the lead in the [8 8 French 8] soccer league on Wednesday . [9 9 Smicer 9] pushed the ball home in injury time to lead his team to a 3-2 victory over [10 10 Montpellier 10] , who were leading 2-1 until [11 11 Cameroon 11] 's [12 12 Marc-Vivien Foe 12] equalised on a header in the 85th minute . The win was the fourth in as many matches this season for [13 13 Lens 13] , who lead the table on 12 points . In-form [14 14 Paris St Germain 14] , who dismissed [15 15 Nantes 15] 1-0 , are second with 10 points . Along with [16 16 Smicer 16] , [17 17 Robert Pires 17] was the star of the night in [18 18 France 18] , scoring the first hat-trick of the league season in [19 19 Metz 19] 's 3-1 home victory over neighbouring [20 20 Strasbourg 20] . [21 21 Pires 21] , one of the most promising strikers in the country , was called up for the first time this week by [22 22 French 22] manager [23 23 Aime Jacquet 23] for a friendly against [24 24 Mexico 24] on Saturday at the [25 25 Parc des Princes 25] . [26 26 Pires 26] scored first with a powerful shot in the 35th minute before striking again from close range just before the break . A solitary raid allowed him to score his third in the 74th . [27 27 Smicer 27] 's goal was as hard-won as his team 's victory . Spurred by Foe 's leveller five minutes before , [28 28 Lens 28] pressed hard and Foe hit the crossbar in the dying seconds on another header . The ball bounced back to [29 29 Smicer 29] 's feet and he scored . [30 30 Montpellier 30] seized an unexpected lead thanks to [31 31 Kader Ferhaoui 31] in the fourth minute after a blunder from [32 32 Lens 32] goalkeeper Jean-Claude Nadon . The side from northern [33 33 France 33] , forced to fight an uphill battle from then on , pulled level thanks to [34 34 Tony Vairelles 34] in the eighth minute but young striker Fabien Lefevre made it two for [35 35 Montpellier 35] five minutes later . League favourites [36 36 PSG 36] scored a convincing 1-0 win over [37 37 Nantes 37] and confirmed they would again be the team to beat this season . Ironically , [38 38 PSG 38] 's victory owed a lot to two former [39 39 Nantes 39] players , striker [40 40 Patrice Loko 40] , who scored on a brilliant shot in the 33rd minute , and defender [41 41 Benoit Cauet 41] , who started the one-two which allowed [42 42 Loko 42] to score . The Parisians , who have yet to concede a goal , were without [43 43 Brazil 43] 's Leonardo and [44 44 Panama 44] 's [45 45 Julio Cesar Dely Valdes 45] , both called up by their national sides . For [46 46 Nantes 46] , who shocked [47 47 PSG 47] to win the league crown two years ago , the fall is very painful . The

Canaries , who lost most of their key players within two years , have yet to win a match this season . Reigning champions [48 48 Auxerre 48] had to settle for a goalless draw against [49 49 Marseille 49] on Tuesday .

- 1: 549482 Vladimír_Šmicer
- 2: 1075697 RC_Lens
- 3: 22989 Paris
- 4: 555876 UEFA_Euro_1996
- 5: 549482 Vladimír_Šmicer
- 6: 5321 Czech_Republic
- 7: 1075697 RC_Lens
- 8: 5843419 France
- 9: 549482 Vladimír_Šmicer
- 10: 1640865 Montpellier_HSC
- 11: 5447 Cameroon
- 12: 253764 Marc-Vivien_Foé
- 13: 1075697 RC_Lens
- 14: 357488 Paris_Saint-Germain_F.C.
- 15: 1135276 FC_Nantes
- 16: 549482 Vladimír_Šmicer
- 17: 553772 Robert_Pirès
- 18: 5843419 France
- 19: 1135310 FC_Metz
- 20: 1256446 RC_Strasbourg
- 21: 553772 Robert_Pirès
- 22: 250197 France_national_football_team
- 23: 2937139 Aimé_Jacquet
- 24: 808402 Mexico_national_football_team
- 25: 761302 Parc_des_Princes
- 26: 553772 Robert_Pirès
- 27: 549482 Vladimír_Šmicer
- 28: 1075697 RC_Lens
- 29: 549482 Vladimír_Šmicer
- 30: 1640865 Montpellier_HSC
- 31: 13121884 Abdelkader_Ferhaoui
- 32: 1075697 RC_Lens
- 33: 5843419 France
- 34: 6026473 Tony_Vairelles
- 35: 1640865 Montpellier_HSC
- 36: 357488 Paris_Saint-Germain_F.C.
- 37: 1135276 FC_Nantes
- 38: 357488 Paris_Saint-Germain_F.C.
- 39: 1135276 FC_Nantes
- 40: 3254011 Patrice_Loko
- 41: 7189794 Benoît_Cauet
- 42: 3254011 Patrice_Loko
- 43: 3383 Brazil
- 44: 22997 Panama
- 45: 5887619 Julio_Dely_Valdés
- 46: 1135276 FC_Nantes
- 47: 357488 Paris_Saint-Germain_F.C.
- 48: 990753 AJ_Auxerre

```
49: 492068 Olympique_de_Marseille
1: 549482 Vladimír Šmicer
    [pem_pos=0, final_scores=0.703192,
final_local_scores=0.7534209, pem=0.0, similarity=26.573296]
2: 1075697 RC_Lens
    [pem_pos=0, final_scores=0.64245135,
final_local_scores=0.61046576, pem=-0.86081105,
similarity=26.988455]
3: 22989 Paris
    [pem_pos=0, final_scores=0.47747415,
final_local_scores=0.53549325, pem=-0.115301244,
similarity=20.696926]
4: 555876 UEFA_Euro_1996
    [pem_pos=0, final_scores=0.4345861,
final_local_scores=0.4303816, pem=0.0, similarity=16.933159]
5: 549482 Vladimír Šmicer
    [pem_pos=0, final_scores=0.8144513,
final_local_scores=0.87495476, pem=0.0, similarity=30.200108]
6: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.3424675,
final_local_scores=0.41411662, pem=-0.27794555,
similarity=17.95929]
7: 1075697 RC_Lens
    [pem_pos=0, final_scores=0.5051399,
final_local_scores=0.4604739, pem=-0.86081105,
similarity=22.512398]
8: 5843419 France
    [pem_pos=0, final_scores=0.66553086,
final_local_scores=0.7311242, pem=-0.80223274,
similarity=30.270584]
9: 549482 Vladimír Šmicer
    [pem_pos=0, final_scores=0.3744297,
final_local_scores=0.39429802, pem=0.0, similarity=15.856353]
10: 1640865 Montpellier_HSC
    [pem_pos=1, final_scores=0.4808492,
final_local_scores=0.44066036, pem=-1.476352,
similarity=25.268528]
11: 5447 Cameroon
    [pem_pos=0, final_scores=0.59484875,
final_local_scores=0.71882427, pem=-0.36813655,
similarity=27.542847]
12: 253764 Marc-Vivien_Foé
    [pem_pos=0, final_scores=0.6287695,
final_local_scores=0.677818, pem=0.0, similarity=24.317154]
13: 1075697 RC_Lens
    [pem_pos=0, final_scores=0.5346634,
final_local_scores=0.49272382, pem=-0.86081105,
similarity=23.4748]
14: 357488 Paris_Saint-Germain_F.C.
    [pem_pos=0, final_scores=0.51581967,
final_local_scores=0.4651931, pem=0.0, similarity=17.972002]
15: 1135276 FC_Nantes
```

[pem_pos=1, final_scores=0.7378728,
final_local_scores=0.7158373, pem=-1.4553026,
similarity=33.365887]
16: 549482 Vladimír Šmicer
[pem_pos=0, final_scores=0.519474,
final_local_scores=0.55273694, pem=0.0, similarity=20.584486]
17: 553772 Robert Pirès
[pem_pos=0, final_scores=0.85008866,
final_local_scores=0.8868205, pem=0.0, similarity=30.554205]
18: 5843419 France
[pem_pos=0, final_scores=0.66131383,
final_local_scores=0.7265178, pem=-0.16240834,
similarity=26.653654]
19: 1135310 FC Metz
[pem_pos=1, final_scores=0.63450235,
final_local_scores=0.6157809, pem=-1.9220247,
similarity=32.918114]
20: 1256446 RC Strasbourg
[pem_pos=1, final_scores=0.5931516,
final_local_scores=0.5729591, pem=-2.2163424,
similarity=33.240772]
21: 553772 Robert Pirès
[pem_pos=0, final_scores=0.7613953,
final_local_scores=0.79394746, pem=0.0, similarity=27.782688]
22: 250197 France_national_football_team
[pem_pos=18, final_scores=0.43910152,
final_local_scores=0.38797957, pem=-5.6158586,
similarity=46.20771]
23: 2937139 Aimé Jacquet
[pem_pos=0, final_scores=0.8463538,
final_local_scores=0.90264994, pem=0.0, similarity=31.026588]
24: 808402 Mexico_national_football_team
[pem_pos=1, final_scores=0.642515,
final_local_scores=0.65871614, pem=-2.8676722,
similarity=39.341965]
25: 761302 Parc_des_Princes
[pem_pos=0, final_scores=0.53431535,
final_local_scores=0.5402428, pem=0.0, similarity=20.211636]
26: 553772 Robert Pirès
[pem_pos=0, final_scores=0.5825836,
final_local_scores=0.5986228, pem=0.0, similarity=21.953812]
27: 549482 Vladimír Šmicer
[pem_pos=0, final_scores=0.5342034,
final_local_scores=0.5688265, pem=0.0, similarity=21.06463]
28: 1075697 RC Lens
[pem_pos=0, final_scores=0.35963494,
final_local_scores=0.30153185, pem=-0.86081105,
similarity=17.769249]
29: 549482 Vladimír Šmicer
[pem_pos=0, final_scores=0.4704309,
final_local_scores=0.4991647, pem=0.0, similarity=18.985783]
30: 1640865 Montpellier_HSC

```
[pem_pos=1, final_scores=0.47984743,
final_local_scores=0.43956608, pem=-1.476352,
similarity=25.235872]
31: 13121884 Abdelkader_Ferhaoui
    [pem_pos=0, final_scores=0.6172463,
final_local_scores=0.7442952, pem=0.0, similarity=26.300966]
32: 1075697 RC_Lens
    [pem_pos=0, final_scores=0.49443692,
final_local_scores=0.4487825, pem=-0.86081105,
similarity=22.163504]
33: 5843419 France
    [pem_pos=0, final_scores=0.6435459,
final_local_scores=0.70710903, pem=-0.16240834,
similarity=26.074457]
34: 6026473 Tony_Vairelles
    [pem_pos=0, final_scores=0.6724548,
final_local_scores=0.74016565, pem=0.0, similarity=26.177732]
35: 1640865 Montpellier_HSC
    [pem_pos=1, final_scores=0.54645985,
final_local_scores=0.51233006, pem=-1.476352,
similarity=27.407295]
36: 357488 Paris_Saint-Germain_F.C.
    [pem_pos=0, final_scores=0.565623,
final_local_scores=0.5195956, pem=-0.4638402,
similarity=22.117918]
37: 1135276 FC_Nantes
    [pem_pos=1, final_scores=0.6807719,
final_local_scores=0.6534633, pem=-1.4553026,
similarity=31.504522]
38: 357488 Paris_Saint-Germain_F.C.
    [pem_pos=0, final_scores=0.52145743,
final_local_scores=0.4713515, pem=-0.4638402,
similarity=20.678217]
39: 1135276 FC_Nantes
    [pem_pos=1, final_scores=0.6695034,
final_local_scores=0.6411541, pem=-1.4553026,
similarity=31.137192]
40: 3254011 Patrice_Loko
    [pem_pos=0, final_scores=0.5887331,
final_local_scores=0.69009656, pem=0.0, similarity=24.68357]
41: 7189794 Benoît_Cauet
    [pem_pos=0, final_scores=0.6454781,
final_local_scores=0.7431121, pem=0.0, similarity=26.26566]
42: 3254011 Patrice_Loko
    [pem_pos=0, final_scores=0.658048,
final_local_scores=0.7728374, pem=0.0, similarity=27.152723]
43: 3383 Brazil
    [pem_pos=0, final_scores=0.56515205,
final_local_scores=0.66998285, pem=-0.21313342,
similarity=25.24239]
44: 22997 Panama
    [pem_pos=0, final_scores=0.41810942,
```

```
final_local_scores=0.54668236, pem=-0.32632124,  
similarity=22.17839]  
45: 5887619 Julio_Dely_Valdés  
    [pem_pos=0, final_scores=0.4399541,  
final_local_scores=0.5027329, pem=0.0, similarity=19.092264]  
46: 1135276 FC_Nantes  
    [pem_pos=1, final_scores=0.59969467,  
final_local_scores=0.56489867, pem=-1.4553026,  
similarity=28.861578]  
47: 357488 Paris_Saint-Germain_F.C.  
    [pem_pos=0, final_scores=0.51528263,  
final_local_scores=0.46460646, pem=-0.4638402,  
similarity=20.476933]  
48: 990753 AJ_Auxerre  
    [pem_pos=0, final_scores=0.60060424,  
final_local_scores=0.5704497, pem=-0.30518255,  
similarity=22.772701]  
49: 492068 Olympique_de_Marseille  
    [pem_pos=1, final_scores=0.54991096,  
final_local_scores=0.5022138, pem=-1.4345005,  
similarity=26.87781]
```

Sample: b'679_SOCCER'

SOCCER - [1 1 DUTCH 1] FIRST DIVISION SUMMARIES . [2 2 AMSTE RDAM 2] 1996-08-28 Summary of [3 3 Dutch 3] first division soccer played on Wednesday : [4 4 Willem II Tilburg 4] 1 (Van Hintum 69th penalty) RKC Waalijk 2 ([5 5 Schreuder 5] 39th , Van Arum 76th , 83rd) . Halftime 0-1 . Attendance 6,150 . [6 6 Vitesse Arnhem 6] 1 ([7 7 Vierklau 7] 85th) [8 8 Sparta Rotterdam 8] 1 ([9 9 Gerard de Nooijer 9] 80th) . Halftime 0-0 . Attendance 5,696 . [10 10 Utrecht 10] 0 [11 11 Twente Enschede 11] 0 . Attendance 9,000 . [12 12 Groningen 12] 1 (Gorre 66th) [13 13 Roda JC Kerkrade 13] 1 ([14 14 Vuren 14] 3rd) Halftime 0-1 . Attendance 10,000 . [15 15 Feyenoord 15] 2 ([16 16 Larsson 16] 73rd , [17 17 Van Gastel 17] 83rd penalty) Graafschap Doetinchem 1 (Schultz 4th) . Halftime 0-1 . Attendance 22,434 . [18 18 Volendam 18] 1 (Jongsma 85th) [19 19 PSV Eindhoven 19] 3 ([20 20 Marcelo 20] 19th , 47th , 56rd) . Halftime 0-1 . Attendance 6,000 . [21 21 Ajax Amsterdam 21] 1 ([22 22 Frank de Boer 22] 30th) [23 23 AZ Alkmaar 23] 0 . Halftime 1-0 . Attendance 48,123 . Played on Tuesday . [24 24 Fortuna Sittard 24] 2 (1 Jeffrey 7th , Roest 33rd) [25 25 Heerenveen 25] 4 ([26 26 Korneev 26] 15th , Hansma 24th , Woudenberg 70th , 90th) . Halftime 2-2 . Attendance 4,000 .

1: 21148 Netherlands
2: 844 Amsterdam
3: 21148 Netherlands
4: 834275 Willem_II_(football_club)
5: 24901774 Dick_Schreuder
6: 834256 Vitesse
7: 9712886 Ferdi_Vierklau

```
8: 79818 Sparta_Rotterdam
9: 9343903 Gérard_de_Nooijer
10: 834226 FC_Utrecht
11: 800033 FC_Twente
12: 1103112 FC_Groningen
13: 834196 Roda_JC_Kerkrade
14: 23592721 Edwin_Vurens
15: 47707 Feyenoord
16: 240952 Henrik_Larsson
17: 9420542 Jean-Paul_van_Gastel
18: 2202202 FC_Volendam
19: 79820 PSV_Eindhoven
20: 16001137 Marcelo_Silva_Ramos
21: 2273 AFC_Ajax
22: 685945 Frank_de_Boer
23: 2289 AZ_Alkmaar
24: 1336325 Fortuna_Sittard
25: 745757 SC_Heerenveen
26: 3989180 Igor_Korneev
1: 5109955 Michael_Jeffrey
    [winner: 40087376 Jared_Jeffrey, pem_pos=13,
final_scores=-0.32371157, final_local_scores=-0.31188935,
pem=-4.1037445, similarity=17.099113]
    RECALL MISS: 5109955 Michael_Jeffrey]
1: 21148 Netherlands
    [pem_pos=0, final_scores=0.62840044,
final_local_scores=0.6797539, pem=-0.7943452,
similarity=28.694698]
2: 844 Amsterdam
    [pem_pos=0, final_scores=0.57511574,
final_local_scores=0.63000417, pem=-0.0642803,
similarity=23.23986]
3: 21148 Netherlands
    [pem_pos=0, final_scores=0.40010065,
final_local_scores=0.43037105, pem=-0.7943452,
similarity=21.252617]
4: 834275 Willem_II_(football_club)
    [pem_pos=0, final_scores=0.77589077,
final_local_scores=0.75228685, pem=0.0, similarity=26.539454]
5: 24901774 Dick_Schreuder
    [pem_pos=3, final_scores=0.44230223,
final_local_scores=0.50329226, pem=-1.609438,
similarity=27.861332]
6: 834256 Vitesse
    [pem_pos=0, final_scores=0.56737465,
final_local_scores=0.5286712, pem=0.0, similarity=19.866316]
7: 9712886 Ferdi_Vierklau
    [pem_pos=0, final_scores=0.53019285,
final_local_scores=0.60916543, pem=0.0, similarity=22.268425]
8: 79818 Sparta_Rotterdam
    [pem_pos=0, final_scores=0.4978004,
final_local_scores=0.46029687, pem=0.0, similarity=17.82589]
```

9: 9343903 Gérard_de_Nooijer
[pem_pos=0, final_scores=0.51464975,
final_local_scores=0.61552435, pem=0.0, similarity=22.458187]
10: 834226 FC_Utrecht
[pem_pos=2, final_scores=0.74635786,
final_local_scores=0.7252937, pem=-2.2398934,
similarity=37.91481]
11: 800033 FC_Twente
[pem_pos=0, final_scores=0.58712566,
final_local_scores=0.5468584, pem=0.0, similarity=20.409058]
12: 1103112 FC_Groningen
[pem_pos=2, final_scores=0.7188282,
final_local_scores=0.6865894, pem=-1.9677447,
similarity=35.27981]
13: 834196 Roda_JC_Kerkrade
[pem_pos=0, final_scores=0.6891715,
final_local_scores=0.6704911, pem=0.0, similarity=24.098505]
14: 23592721 Edwin_Vurens
[pem_pos=0, final_scores=0.5709067,
final_local_scores=0.63268083, pem=0.0, similarity=22.970171]
15: 47707 Feyenoord
[pem_pos=0, final_scores=0.58612674,
final_local_scores=0.53916097, pem=0.0, similarity=20.179352]
16: 240952 Henrik_Larsson
[pem_pos=0, final_scores=0.42707402,
final_local_scores=0.448573, pem=-0.7646139,
similarity=21.634115]
17: 9420542 Jean-Paul_van_Gastel
[pem_pos=0, final_scores=0.5541566,
final_local_scores=0.6251414, pem=0.0, similarity=22.745178]
18: 2202202 FC_Volendam
[pem_pos=1, final_scores=0.6125242,
final_local_scores=0.5898991, pem=-0.8958714,
similarity=26.565367]
19: 79820 PSV_Eindhoven
[pem_pos=0, final_scores=0.531552,
final_local_scores=0.4808377, pem=0.0, similarity=18.43887]
20: 16001137 Marcelo_Silva_Ramos
[pem_pos=13, final_scores=0.5816437,
final_local_scores=0.7603566, pem=-3.61824,
similarity=46.45682]
21: 2273 AFC_Ajax
[pem_pos=0, final_scores=0.62524265,
final_local_scores=0.5873083, pem=0.0, similarity=21.616163]
22: 685945 Frank_de_Boer
[pem_pos=0, final_scores=0.75228566,
final_local_scores=0.7880263, pem=0.0, similarity=27.60599]
23: 2289 AZ_Alkmaar
[pem_pos=0, final_scores=0.44518584,
final_local_scores=0.3891015, pem=0.0, similarity=15.701279]
24: 1336325 Fortuna_Sittard
[pem_pos=0, final_scores=0.73605084,

```
final_local_scores=0.7220226, pem=0.0, similarity=25.636307]
25: 745757 SC_Heerenveen
    [pem_pos=1, final_scores=0.6559785,
final_local_scores=0.62361914, pem=-0.6351416,
similarity=26.15375]
26: 3989180 Igor_Korneev
    [pem_pos=0, final_scores=0.53177863,
final_local_scores=0.62680936, pem=0.0, similarity=22.794954]
```

Sample: b'680_SOCCER'

SOCCER - [1 1 GERMAN 1] FIRST DIVISION RESULTS / STANDINGS .
[2 2 BONN 2] 1996-08-28 Results of [3 3 German 3] first
division soccer matches on Wednesday : [4 4 Karlsruhe 4] 4 [5
5 St Pauli 5] 0 [6 6 Bayern Munich 6] 4 [7 7 Bayer Leverkusen
7] 2 [8 8 Cologne 8] 0 [9 9 Hansa Rostock 9] 2 [10 10 Fortuna
Duesseldorf 10] 0 [11 11 1860 Munich 11] 0 [12 12 Arminia Bie
lefeld 12] 1 [13 13 Duisburg 13] 1 Standings (tabulated
under played , won , drawn , lost , goals for , against ,
points) : [14 14 Bayern Munich 14] 4 3 1 0 11 4 10 [15 15 VfB
Stuttgart 15] 3 3 0 0 10 1 9 [16 16 Borussia Dortmund 16] 4 3
0 1 12 6 9 [17 17 Cologne 17] 4 3 0 1 7 3 9 [18 18 Karlsruhe 1
8] 3 2 1 0 9 3 7 [19 19 Bayer Leverkusen 19] 4 2 0 2 9 8 6 [20
20 VfL Bochum 20] 4 1 3 0 4 3 6 [21 21 SV Hamburg 21] 4 2 0 2
7 7 6 [22 22 Hansa Rostock 22] 4 1 2 1 5 4 5 [23 23 Werder Br
emen 23] 4 1 1 2 5 6 4 [24 24 Munich 1860 24] 4 1 1 2 3 5 4 [2
5 25 St Pauli 25] 4 1 1 2 7 11 4 [26 26 Fortuna Duesseldorf 26
] 4 1 1 2 1 7 4 [27 27 Arminia Bielefeld 27] 4 0 3 1 3 4 3 [28
28 Schalke 04 28] 4 0 3 1 5 9 3 [29 29 Freiburg 29] 4 1 0 3 6
13 3 [30 30 Borussia Moenchengladbach 30] 4 0 2 2 1 4 2 [31 31
Duisburg 31] 4 0 1 3 2 9 1

1: 11867 Germany

2: 3295 Bonn

3: 11867 Germany

4: 1734592 Karlsruher_SC

5: 881244 FC_St._Pauli

6: 172326 FC_Bayern_Munich

7: 502005 Bayer_04_Leverkusen

8: 497888 1._FC_Koeln

9: 599978 F.C._Hansa_Rostock

10: 1949688 Fortuna_Duesseldorf

11: 1007441 TSV_1860_Muenchen

12: 1145228 Arminia_Bielefeld

13: 1586579 MSV_Duisburg

14: 172326 FC_Bayern_Munich

15: 737608 VfB_Stuttgart

16: 331715 Borussia_Dortmund

17: 497888 1._FC_Koeln

18: 1734592 Karlsruher_SC

19: 502005 Bayer_04_Leverkusen

20: 513623 VfL_Bochum

21: 535879 Hamburger_SV

22: 599978 F.C._Hansa_Rostock

```
23: 7598082 SV_Werder_Bremen
24: 1007441 TSV_1860_München
25: 881244 FC_St._Pauli
26: 1949688 Fortuna_Düsseldorf
27: 1145228 Arminia_Bielefeld
28: 410348 FC_Schalke_04
29: 1438835 SC_Freiburg
30: 263466 Borussia_Mönchengladbach
31: 1586579 MSV_Duisburg
1: 11867 Germany
    [pem_pos=1, final_scores=0.78010327,
final_local_scores=0.8649227, pem=-1.0465823,
similarity=35.59221]
2: 3295 Bonn
    [pem_pos=0, final_scores=0.3877489,
final_local_scores=0.459961, pem=0.0, similarity=17.815866]
3: 11867 Germany
    [pem_pos=1, final_scores=0.7537375,
final_local_scores=0.8361221, pem=-1.0465823,
similarity=34.732742]
4: 1734592 Karlsruher_SC
    [pem_pos=1, final_scores=0.7181671,
final_local_scores=0.68306834, pem=-2.2263577,
similarity=36.581116]
5: 881244 FC_St._Pauli
    [pem_pos=1, final_scores=0.5392741,
final_local_scores=0.49891657, pem=-0.11686432,
similarity=19.613905]
6: 172326 FC_Bayern_Munich
    [pem_pos=0, final_scores=0.48282605,
final_local_scores=0.427046, pem=0.0, similarity=16.833618]
7: 502005 Bayer_04_Leverkusen
    [pem_pos=0, final_scores=0.52444696,
final_local_scores=0.46752572, pem=0.0, similarity=18.041613]
8: 497888 1._FC_Köln
    [pem_pos=10, final_scores=0.8019739,
final_local_scores=0.76833653, pem=-3.09805,
similarity=43.866085]
9: 599978 F.C._Hansa_Rostock
    [pem_pos=0, final_scores=0.59588045,
final_local_scores=0.5570722, pem=0.0, similarity=20.71386]
10: 1949688 Fortuna_Düsseldorf
    [pem_pos=0, final_scores=0.6153402,
final_local_scores=0.5740108, pem=0.0, similarity=21.21934]
11: 1007441 TSV_1860_München
    [pem_pos=0, final_scores=0.45279342,
final_local_scores=0.39667696, pem=0.0, similarity=15.927345]
12: 1145228 Arminia_Bielefeld
    [pem_pos=0, final_scores=0.48938984,
final_local_scores=0.4358608, pem=0.0, similarity=17.096668]
13: 1586579 MSV_Duisburg
    [pem_pos=1, final_scores=0.46620774,
```

```
final_local_scores=0.41397238, pem=-1.5900439,  
similarity=25.09038]  
14: 172326 FC_Bayern_Munich  
    [pem_pos=0, final_scores=0.44879228,  
final_local_scores=0.38986933, pem=0.0, similarity=15.724192]  
15: 737608 VfB_Stuttgart  
    [pem_pos=0, final_scores=0.44830197,  
final_local_scores=0.3882141, pem=0.0, similarity=15.674796]  
16: 331715 Borussia_Dortmund  
    [pem_pos=0, final_scores=0.48460162,  
final_local_scores=0.42613733, pem=0.0, similarity=16.806501]  
17: 497888 1._FC_Köln  
    [pem_pos=10, final_scores=0.75546354,  
final_local_scores=0.71753097, pem=-3.09805,  
similarity=42.349945]  
18: 1734592 Karlsruher_SC  
    [pem_pos=1, final_scores=0.685021,  
final_local_scores=0.64686126, pem=-2.2263577,  
similarity=35.50062]  
19: 502005 Bayer_04_Leverkusen  
    [pem_pos=0, final_scores=0.5455907,  
final_local_scores=0.49062204, pem=0.0, similarity=18.730854]  
20: 513623 VfL_Bochum  
    [pem_pos=0, final_scores=0.5693969,  
final_local_scores=0.5251786, pem=0.0, similarity=19.76209]  
21: 535879 Hamburger_SV  
    [pem_pos=0, final_scores=0.62342906,  
final_local_scores=0.5753285, pem=0.0, similarity=21.258663]  
22: 599978 F.C._Hansa_Rostock  
    [pem_pos=0, final_scores=0.5588675,  
final_local_scores=0.5166412, pem=0.0, similarity=19.507317]  
23: 7598082 SV_Werder_Bremen  
    [pem_pos=0, final_scores=0.6127834,  
final_local_scores=0.56802106, pem=0.0, similarity=21.040594]  
24: 1007441 TSV_1860_München  
    [pem_pos=0, final_scores=0.49297,  
final_local_scores=0.44056374, pem=0.0, similarity=17.237013]  
25: 881244 FC_St._Pauli  
    [pem_pos=1, final_scores=0.5778504,  
final_local_scores=0.54105526, pem=-0.11686432,  
similarity=20.871408]  
26: 1949688 Fortuna_Düsseldorf  
    [pem_pos=0, final_scores=0.65984803,  
final_local_scores=0.6226289, pem=0.0, similarity=22.670202]  
27: 1145228 Arminia_Bielefeld  
    [pem_pos=0, final_scores=0.51216537,  
final_local_scores=0.4607396, pem=0.0, similarity=17.839102]  
28: 410348 FC_Schalke_04  
    [pem_pos=0, final_scores=0.48822433,  
final_local_scores=0.42789626, pem=0.0, similarity=16.858992]  
29: 1438835 SC_Freiburg  
    [pem_pos=2, final_scores=0.8354512,
```

```
final_local_scores=0.83067995, pem=-1.7450409,  
similarity=38.368664]  
30: 263466 Borussia_Mönchengladbach  
    [pem_pos=0, final_scores=0.5165463,  
final_local_scores=0.46086973, pem=0.0, similarity=17.842985]  
31: 1586579 MSV_Duisburg  
    [pem_pos=1, final_scores=0.45888525,  
final_local_scores=0.4059736, pem=-1.5900439,  
similarity=24.85168]
```

Sample: b'681_SOCCER'

```
SOCCER - [1 1 FRENCH 1] LEAGUE SUMMARIES . [2 2 PARIS 2]  
1996-08-28 Summaries of [3 3 French 3] first division matches  
on Wednesday : [4 4 Bastia 4] 0 [5 5 Lille 5] 0 . 0-0 .  
5,000 . [6 6 Cannes 6] 0 [7 7 Monaco 7] 2 ( [8 8 Henry 8]  
26th , 71st ) . 0-1 . 7,000 . [9 9 Le Havre 9] 1 ( Samson  
24th ) [10 10 Caen 10] 1 ( Etienne Mendy 4th ) . 1-1 .  
12,000 . [11 11 Lens 11] 3 ( [12 12 Vairelles 12] 8th , Foe  
85th , [13 13 Smicer 13] 90th ) [14 14 Montpellier 14] 2 ( [15  
15 Ferhaoui 15] 4th , Lefevre 13th ) . 1-2 . 30,000 . [16 16 L  
yon 16] 2 ( [17 17 Caveglia 17] 23rd , [18 18 Giuly 18]  
30th ) [19 19 Nancy 19] 0 . 2-0 . 15,000 . [20 20 Metz 20] 3  
( [21 21 Pires 21] 35th , 48th , 74th ) [22 22 Strasbourg 22]  
1 ( 1 Rodriguez 56th ) . 1-0 . 14,000 . [23 23 Nice 23] 1  
( Chaouch 64th ) [24 24 Guingamp 24] 2 ( Rouxel 10th , Baret  
89th ) . 0-1 . 4,000 . [25 25 Paris St Germain 25] 1 ( [26 26  
Loko 26] 33rd ) [27 27 Nantes 27] 0 . 1-0 . 30,000 . [28 28 Re  
nnes 28] 1 ( Guiavarch 27th ) [29 29 Bordeaux 29] 1 ( [30 30 Co  
lleter 30] 86th ) . 1-0 . 16,000 .
```

```
1: 5843419 France  
2: 22989 Paris  
3: 5843419 France  
4: 1256442 SC_Bastia  
5: 990757 Lille_OSC  
6: 2237150 AS_Cannes  
7: 222981 AS_Monaco_FC  
8: 220160 Thierry_Henry  
9: 3658991 Le_Havre_AC  
10: 1256445 Stade_Malherbe_Caen  
11: 1075697 RC_Lens  
12: 6026473 Tony_Vairelles  
13: 549482 Vladimír_Smicer  
14: 1640865 Montpellier_HSC  
15: 13121884 Abdelkader_Ferhaoui  
16: 979521 Olympique_Lyonnais  
17: 11576169 Alain_Caveglia  
18: 1706459 Ludovic_Giuly  
19: 1640855 AS_Nancy  
20: 1135310 FC_Metz  
21: 553772 Robert_Pirès  
22: 1256446 RC_Strasbourg  
23: 704478 OGC_Nice
```

```
24: 1693070 En_Avant_de_Guingamp
25: 357488 Paris_Saint-Germain_F.C.
26: 3254011 Patrice_Loko
27: 1135276 FC_Nantes
28: 1004187 Stade_Rennais_F.C.
29: 1135285 FC_Girondins_de_Bordeaux
30: 7737628 Patrick_Colleter
1: 7213440 Bruno_Rodriguez
    [winner: 2385554 Julien_Rodriguez, pem_pos=10,
final_scores=-0.22966337, final_local_scores=-0.2373994,
pem=-3.4915981, similarity=15.993094]
    RECALL MISS: 7213440 Bruno_Rodriguez]
1: 5843419 France
    [pem_pos=0, final_scores=0.6563052,
final_local_scores=0.7315398, pem=-0.80223274,
similarity=30.282986]
2: 22989 Paris
    [pem_pos=0, final_scores=0.5112838,
final_local_scores=0.5761495, pem=-0.115301244,
similarity=21.91019]
3: 5843419 France
    [pem_pos=0, final_scores=0.70197535,
final_local_scores=0.78142744, pem=-0.80223274,
similarity=31.771734]
4: 1256442 SC_Bastia
    [pem_pos=1, final_scores=0.5599173,
final_local_scores=0.522937, pem=-0.7022244,
similarity=23.514]
5: 990757 Lille OSC
    [pem_pos=1, final_scores=0.6996582,
final_local_scores=0.6636229, pem=-1.3501564,
similarity=31.235903]
6: 2237150 AS_Cannes
    [pem_pos=2, final_scores=0.6590806,
final_local_scores=0.643547, pem=-2.2011297,
similarity=35.264523]
7: 222981 AS_Monaco_FC
    [pem_pos=1, final_scores=0.9217127,
final_local_scores=0.90827817, pem=-2.451784,
similarity=44.52773]
8: 220160 Thierry_Henry
    [pem_pos=0, final_scores=0.7416985,
final_local_scores=0.78081155, pem=-2.3426633,
similarity=40.130455]
9: 3658991 Le_Havre_AC
    [pem_pos=1, final_scores=0.49670416,
final_local_scores=0.45460898, pem=-1.3456037,
similarity=24.973753]
10: 1256445 Stade_Malherbe_Caen
    [pem_pos=1, final_scores=0.5068088,
final_local_scores=0.46625566, pem=-1.6168041,
similarity=26.796144]
```

```
11: 1075697 RC_Lens
    [pem_pos=0, final_scores=0.51581746,
final_local_scores=0.46573895, pem=-0.86081105,
similarity=22.669518]
12: 6026473 Tony_Vairelles
    [pem_pos=0, final_scores=0.39762056,
final_local_scores=0.43144226, pem=0.0, similarity=16.964811]
13: 549482 Vladimír_Šmicer
    [pem_pos=0, final_scores=0.39501536,
final_local_scores=0.4519834, pem=0.0, similarity=17.577799]
14: 1640865 Montpellier_HSC
    [pem_pos=1, final_scores=0.57648736,
final_local_scores=0.53671265, pem=-1.476352,
similarity=28.13492]
15: 13121884 Abdelkader_Ferhaoui
    [pem_pos=0, final_scores=0.5726586,
final_local_scores=0.68571705, pem=0.0, similarity=24.552877]
16: 979521 Olympique_Lyonnais
    [pem_pos=1, final_scores=0.71505314,
final_local_scores=0.67844266, pem=-1.9193003,
similarity=34.773247]
17: 11576169 Alain_Caveglia
    [pem_pos=0, final_scores=0.5256189,
final_local_scores=0.59596276, pem=0.0, similarity=21.87443]
18: 1706459 Ludovic_Giuly
    [pem_pos=0, final_scores=0.39345074,
final_local_scores=0.38388133, pem=0.0, similarity=15.545497]
19: 1640855 AS_Nancy
    [pem_pos=0, final_scores=0.7269024,
final_local_scores=0.70002717, pem=-1.6882259,
similarity=34.160755]
20: 1135310 FC_Metz
    [pem_pos=1, final_scores=0.6970833,
final_local_scores=0.6703983, pem=-1.9220247,
similarity=34.548004]
21: 553772 Robert_Pirès
    [pem_pos=0, final_scores=0.4503669,
final_local_scores=0.47057003, pem=-0.5653512,
similarity=21.206928]
22: 1256446 RC_Strasbourg
    [pem_pos=1, final_scores=0.6876985,
final_local_scores=0.6626863, pem=-2.2163424,
similarity=35.91841]
23: 704478 OGC_Nice
    [pem_pos=1, final_scores=0.74989,
final_local_scores=0.72607493, pem=-2.1390686,
similarity=37.389824]
24: 1693070 En_Avant_de_Guingamp
    [pem_pos=0, final_scores=0.47216254,
final_local_scores=0.43766636, pem=0.0, similarity=17.15055]
25: 357488 Paris_Saint-Germain_F.C.
    [pem_pos=0, final_scores=0.5007661,
```

```

final_local_scores=0.44469512, pem=0.0, similarity=17.360302]
26: 3254011 Patrice_Loko
    [pem_pos=0, final_scores=0.37811333,
final_local_scores=0.47624356, pem=-0.52757347,
similarity=21.170797]
27: 1135276 FC_Nantes
    [pem_pos=1, final_scores=0.74389076,
final_local_scores=0.71590775, pem=-1.4553026,
similarity=33.36799]
28: 1004187 Stade_Rennais_F.C.
    [pem_pos=1, final_scores=0.38981658,
final_local_scores=0.32863647, pem=-1.0138907,
similarity=19.410576]
29: 1135285 FC_Girondins_de_Bordeaux
    [pem_pos=1, final_scores=0.7294833,
final_local_scores=0.69326144, pem=-1.6081072,
similarity=33.523155]
30: 7737628 Patrick_Colleter
    [pem_pos=0, final_scores=0.42312807,
final_local_scores=0.49383658, pem=0.0, similarity=18.826782]

```

Sample: b'682_SOCCER'

```

SOCCER - [1 1 FRENCH 1] LEAGUE STANDINGS . [2 2 PARIS 2]
1996-08-28 Standings in the [3 3 French 3] first division
after Wednesday 's matches ( tabulate under played , won ,
drawn , lost , goals for , against , points ) : [4 4 Lens 4]
4 4 0 0 9 3 12 [5 5 Paris Saint-Germain 5] 4 3 1 0 4 0 10 [6 6
Bastia 6] 4 2 2 0 4 1 8 [7 7 Auxerre 7] 4 2 2 0 3 0 8 [8 8 Mo
naco 8] 4 2 1 1 7 4 7 [9 9 Lyon 9] 4 2 1 1 6 4 7 [10 10 Metz 1
0] 4 2 1 1 6 4 7 [11 11 Lille 11] 4 2 1 1 4 3 7 [12 12 Guinga
mp 12] 4 2 1 1 4 3 7 [13 13 Cannes 13] 4 2 1 1 4 4 7 [14 14 Bo
rdeaux 14] 4 1 3 0 3 2 6 [15 15 Marseille 15] 4 1 2 1 5 4 5 [1
6 16 Rennes 16] 4 1 1 2 5 7 4 [17 17 Strasbourg 17] 4 1 0 3 2
7 3 [18 18 Montpellier 18] 4 0 2 2 3 5 2 [19 19 Le Havre 19]
4 0 2 2 2 4 2 [20 20 Caen 20] 4 0 2 2 2 6 2 [21 21 Nice 21] 4
0 1 3 3 7 1 [22 22 Nantes 22] 4 0 1 3 2 6 1 [23 23 Nancy 23]
4 0 1 3 2 7 1

```

- 1: 5843419 France
- 2: 22989 Paris
- 3: 5843419 France
- 4: 1075697 RC_Lens
- 5: 357488 Paris_Saint-Germain_F.C.
- 6: 1256442 SC_Bastia
- 7: 990753 AJ_Auxerre
- 8: 222981 AS_Monaco_FC
- 9: 979521 Olympique_Lyonnais
- 10: 1135310 FC_Metz
- 11: 990757 Lille OSC
- 12: 1693070 En_Avant_de_Guingamp
- 13: 2237150 AS_Cannes
- 14: 1135285 FC_Girondins_de_Bordeaux
- 15: 492068 Olympique_de_Marseille

```
16: 1004187 Stade_Rennais_F.C.
17: 1256446 RC_Strasbourg
18: 1640865 Montpellier_HSC
19: 3658991 Le_Havre_AC
20: 1256445 Stade_Malherbe_Caen
21: 704478 OGC_Nice
22: 1135276 FC_Nantes
23: 1640855 AS_Nancy
1: 5843419 France
    [pem_pos=0, final_scores=0.6550791,
final_local_scores=0.72567993, pem=-0.80223274,
similarity=30.108116]
2: 22989 Paris
    [pem_pos=0, final_scores=0.5179841,
final_local_scores=0.57950157, pem=-0.115301244,
similarity=22.010221]
3: 5843419 France
    [pem_pos=0, final_scores=0.58864385,
final_local_scores=0.6531095, pem=-0.80223274,
similarity=27.942469]
4: 1075697 RC_Lens
    [pem_pos=0, final_scores=0.5093902,
final_local_scores=0.45701814, pem=-0.86081105,
similarity=22.409271]
5: 357488 Paris_Saint-Germain_F.C.
    [pem_pos=0, final_scores=0.48491305,
final_local_scores=0.4268896, pem=0.0, similarity=16.82895]
6: 1256442 SC_Bastia
    [pem_pos=1, final_scores=0.44975358,
final_local_scores=0.4002853, pem=-0.7022244,
similarity=19.85383]
7: 990753 AJ_Auxerre
    [pem_pos=0, final_scores=0.6194137,
final_local_scores=0.57802415, pem=-0.30518255,
similarity=22.998737]
8: 222981 AS_Monaco_FC
    [pem_pos=1, final_scores=0.7913497,
final_local_scores=0.76607925, pem=-2.451784,
similarity=40.28423]
9: 979521 Olympique_Lyonnais
    [pem_pos=1, final_scores=0.66605926,
final_local_scores=0.62714785, pem=-1.9193003,
similarity=33.242508]
10: 1135310 FC_Metz
    [pem_pos=1, final_scores=0.7375094,
final_local_scores=0.7095247, pem=-1.9220247,
similarity=35.715614]
11: 990757 Lille OSC
    [pem_pos=1, final_scores=0.63090795,
final_local_scores=0.5880156, pem=-1.3501564,
similarity=28.979631]
12: 1693070 En_Avant_de_Guingamp
```

```
[pem_pos=0, final_scores=0.4771585,
final_local_scores=0.44003075, pem=0.0, similarity=17.221107]
13: 2237150 AS_Cannes
    [pem_pos=2, final_scores=0.62587315,
final_local_scores=0.6022164, pem=-2.2011297,
similarity=34.031136]
14: 1135285 FC_Girondins_de_Bordeaux
    [pem_pos=1, final_scores=0.55774295,
final_local_scores=0.5044088, pem=-1.6081072,
similarity=27.887415]
15: 492068 Olympique_de_Marseille
    [pem_pos=1, final_scores=0.6019798,
final_local_scores=0.5494284, pem=-1.4345005,
similarity=28.286789]
16: 1004187 Stade_Rennais_F.C.
    [pem_pos=1, final_scores=0.38964683,
final_local_scores=0.32648146, pem=-1.0138907,
similarity=19.346266]
17: 1256446 RC_Strasbourg
    [pem_pos=1, final_scores=0.7082049,
final_local_scores=0.6818865, pem=-2.2163424,
similarity=36.491383]
18: 1640865 Montpellier_HSC
    [pem_pos=1, final_scores=0.5487729,
final_local_scores=0.5033212, pem=-1.476352,
similarity=27.138453]
19: 3658991 Le_Havre_AC
    [pem_pos=1, final_scores=0.49401456,
final_local_scores=0.44922036, pem=-1.3456037,
similarity=24.812946]
20: 1256445 Stade_Malherbe_Caen
    [pem_pos=1, final_scores=0.46122104,
final_local_scores=0.41410667, pem=-1.6168041,
similarity=25.239914]
21: 704478 OGC_Nice
    [pem_pos=1, final_scores=0.76869136,
final_local_scores=0.7429391, pem=-2.1390686,
similarity=37.893085]
22: 1135276 FC_Nantes
    [pem_pos=1, final_scores=0.57121927,
final_local_scores=0.5265261, pem=-1.4553026,
similarity=27.716463]
23: 1640855 AS_Nancy
    [pem_pos=0, final_scores=0.66188407,
final_local_scores=0.6284453, pem=-1.6882259,
similarity=32.02461]
```

Sample: b'683_SOCCER'
SOCCER - [1 1 FRENCH 1] LEAGUE RESULTS . [2 2 PARIS 2]
1996-08-28 [3 3 French 3] first division soccer matches on
Wednesday : [4 4 Paris SG 4] 1 [5 5 Nantes 5] 0 [6 6 Lens 6]
3 [7 7 Montpellier 7] 2 [8 8 Bastia 8] 0 [9 9 Lille 9] 0 [10 1

0 Cannes 10] 0 [11 11 Monaco 11] 2 [12 12 Rennes 12] 1 [13 13
Bordeaux 13] 1 [14 14 Lyon 14] 2 [15 15 Nancy 15] 0 [16 16 Ni
ce 16] 1 [17 17 Guingamp 17] 2 [18 18 Metz 18] 3 [19 19 Stras
bourg 19] 1 [20 20 Le Havre 20] 1 [21 21 Caen 21] 1 Played
Tuesday : [22 22 Auxerre 22] 0 [23 23 Marseille 23] 0
1: 5843419 France
2: 22989 Paris
3: 5843419 France
4: 357488 Paris_Saint-Germain_F.C.
5: 1135276 FC_Nantes
6: 1075697 RC_Lens
7: 1640865 Montpellier_HSC
8: 1256442 SC_Bastia
9: 990757 Lille_OSC
10: 2237150 AS_Cannes
11: 222981 AS_Monaco_FC
12: 1004187 Stade_Rennais_F.C.
13: 1135285 FC_Girondins_de_Bordeaux
14: 979521 Olympique_Lyonnais
15: 1640855 AS_Nancy
16: 704478 OGC_Nice
17: 1693070 En_Avant_de_Guingamp
18: 1135310 FC_Metz
19: 1256446 RC_Strasbourg
20: 3658991 Le_Havre_AC
21: 1256445 Stade_Malherbe_Caen
22: 990753 AJ_Auxerre
23: 492068 Olympique_de_Marseille
1: 5843419 France
 [pem_pos=0, final_scores=0.67355144,
final_local_scores=0.74585813, pem=-0.80223274,
similarity=30.710274]
2: 22989 Paris
 [pem_pos=0, final_scores=0.49018306,
final_local_scores=0.5491332, pem=-0.115301244,
similarity=21.103968]
3: 5843419 France
 [pem_pos=0, final_scores=0.7100581,
final_local_scores=0.78573614, pem=-0.80223274,
similarity=31.900312]
4: 357488 Paris_Saint-Germain_F.C.
 [pem_pos=0, final_scores=0.48383492,
final_local_scores=0.42571193, pem=0.0, similarity=16.793806]
5: 1135276 FC_Nantes
 [pem_pos=1, final_scores=0.63446504,
final_local_scores=0.59561247, pem=-1.4553026,
similarity=29.77814]
6: 1075697 RC_Lens
 [pem_pos=0, final_scores=0.52136695,
final_local_scores=0.470101, pem=-0.86081105,
similarity=22.79969]
7: 1640865 Montpellier_HSC

```
[pem_pos=1, final_scores=0.53255403,
final_local_scores=0.48560458, pem=-1.476352,
similarity=26.609753]
8: 1256442 SC_Bastia
    [pem_pos=1, final_scores=0.5069335,
final_local_scores=0.46274573, pem=-0.7022244,
similarity=21.717773]
9: 990757 Lille OSC
    [pem_pos=1, final_scores=0.7276743,
final_local_scores=0.69371814, pem=-1.3501564,
similarity=32.134007]
10: 2237150 AS_Cannes
    [pem_pos=2, final_scores=0.71746206,
final_local_scores=0.7022634, pem=-2.2011297,
similarity=37.01674]
11: 222981 AS_Monaco_FC
    [pem_pos=1, final_scores=0.77831775,
final_local_scores=0.75184375, pem=-2.451784,
similarity=39.859413]
12: 1004187 Stade_Rennais_F.C.
    [pem_pos=1, final_scores=0.418137,
final_local_scores=0.35760272, pem=-1.0138907,
similarity=20.274986]
13: 1135285 FC_Girondins_de_Bordeaux
    [pem_pos=1, final_scores=0.6463992,
final_local_scores=0.60125226, pem=-1.6081072,
similarity=30.777416]
14: 979521 Olympique_Lyonnais
    [pem_pos=1, final_scores=0.69409543,
final_local_scores=0.6577732, pem=-1.9193003,
similarity=34.15643]
15: 1640855 AS_Nancy
    [pem_pos=0, final_scores=0.77026635,
final_local_scores=0.7468365, pem=-1.6882259,
similarity=35.55764]
16: 704478 OGC_Nice
    [pem_pos=1, final_scores=0.7343182,
final_local_scores=0.70539165, pem=-2.1390686,
similarity=36.772594]
17: 1693070 En_Avant_de_Guingamp
    [pem_pos=0, final_scores=0.47083962,
final_local_scores=0.43312836, pem=0.0, similarity=17.015127]
18: 1135310 FC_Metz
    [pem_pos=1, final_scores=0.65278137,
final_local_scores=0.61697215, pem=-1.9220247,
similarity=32.953663]
19: 1256446 RC_Strasbourg
    [pem_pos=1, final_scores=0.7195522,
final_local_scores=0.69428164, pem=-2.2163424,
similarity=36.86128]
20: 3658991 Le_Havre_AC
    [pem_pos=1, final_scores=0.49723327,
```

```
final_local_scores=0.45273632, pem=-1.3456037,
similarity=24.91787]
21: 1256445 Stade_Malherbe_Caen
    [pem_pos=1, final_scores=0.46809238,
final_local_scores=0.42161262, pem=-1.6168041,
similarity=25.463907]
22: 990753 AJ_Auxerre
    [pem_pos=0, final_scores=0.66141343,
final_local_scores=0.6239025, pem=-0.30518255,
similarity=24.367838]
23: 492068 Olympique_de_Marseille
    [pem_pos=1, final_scores=0.5459742,
final_local_scores=0.4882508, pem=-1.4345005,
similarity=26.461126]
```

Sample: b'684 ATHLETICS'

ATHLETICS - [1 1 CHRISTIE 1] AND JOHNSON ASKED TO JOIN [2 2 0 WENS 2] ' TRIBUTE . Adrian Warner [3 3 BONN 3] 1996-08-28 Organisers hope to persuade [4 4 Britain 4] 's former [5 5 0lympic 5] 100 metres champion [6 6 Linford Christie 6] to join a " Dream Team " sprint relay in a special tribute to [7 7 Je sse Owens 7] at Friday 's [8 8 Berlin 8] grand prix . [9 9 Christie 9] , who is retiring from international competition at the end of the season , was not due to compete in the [10 10 German 10] capital but [11 11 Berlin 11] promoter Rudi Thiel said : " We are still hopeful of getting him to come . " Thiel has managed to get most of the [12 12 Olympic 12] 100 metres champions since 1948 to attend the meeting , which is being held in the stadium where [13 13 Owens 13] won four gold medals 60 years ago at the 1 Berlin 2 Olympics . [14 14 Canada 14] 's [15 15 Donovan Bailey 15] , the 3 Olympic 100 metres champion and world record holder , and Namibian [16 16 Frankie Fredericks 16] , the silver medallist at the recent [17 17 Atlanta 17] Games , have already agreed to run in the 4X100 metres team . Thiel said on Wednesday that he had also asked 4 Olympic 200 and 400 champion [18 18 Michael Johnson 18] to run as well as [19 19 Christie 19] . " Most of the [20 20 Olympic 20] champions of the past are coming including [21 21 Britain 21] 's (1980 champion) [22 22 Allan Wells 22] . [23 23 Christie 23] belongs to them . It would be great to have him here . " There is a good offer My minimum would be that he just ran the relay , " he said . The 36-year-old [24 24 Briton 24] is still considering the offer and is expected to announce his decision later on Wednesday . [25 25 Owens 25] 's widow Ruth is not well enough to attend but a message from her will be read out during the meeting and one of the sprinter 's relatives is expected to attend . The relay race , which will include squads from [26 26 Africa 26] , the [27 27 United States 27] and [28 28 Europe 28] as well as the [29 29 Owens 29] ' quartet , will be held at the end of the meeting . Organisers had hoped to include 1984 and 1988 champion [30 30 Carl Lewis 30] in the squad but he injured

himself in [31 31 Brussels 31] last Friday .

1: 67118 Linford_Christie
2: 46827 Jesse_Owens
3: 3295 Bonn
4: 31717 United_Kingdom
5: 27566 Summer_Olympic_Games
6: 67118 Linford_Christie
7: 46827 Jesse_Owens
8: 3354 Berlin
9: 67118 Linford_Christie
10: 11867 Germany
11: 3354 Berlin
12: 27566 Summer_Olympic_Games
13: 46827 Jesse_Owens
14: 5042916 Canada
15: 67116 Donovan_Bailey
16: 71410 Frankie_Fredericks
17: 72311 1996_Summer_Olympics
18: 653621 Michael_Johnson_(sprinter)
19: 67118 Linford_Christie
20: 27566 Summer_Olympic_Games
21: 31717 United_Kingdom
22: 67299 Allan_Wells
23: 67118 Linford_Christie
24: 31717 United_Kingdom
25: 46827 Jesse_Owens
26: 5334607 Africa
27: 3434750 United_States
28: 9239 Europe
29: 46827 Jesse_Owens
30: 67127 Carl_Lewis
31: 3708 Brussels

1: 39721 1936_Summer_Olympics
[winner: 39721 1936_Summer_Olympics, pem_pos=0,
final_scores=-0.16098276, final_local_scores=-0.22430763,
pem=0.0, similarity=-2.604074]
gt: 39721 1936_Summer_Olympics, pem_pos=0,
final_scores=-0.16098276, final_local_scores=-0.22430763,
pem=0.0, similarity=-2.604074]

2: 39721 1936_Summer_Olympics
[winner: 39721 1936_Summer_Olympics, pem_pos=0,
final_scores=-0.16098276, final_local_scores=-0.22430763,
pem=0.0, similarity=-2.604074]
gt: 39721 1936_Summer_Olympics, pem_pos=0,
final_scores=-0.16098276, final_local_scores=-0.22430763,
pem=0.0, similarity=-2.604074]

3: 72311 1996_Summer_Olympics
[winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.25811172, final_local_scores=-0.3341548,
pem=-3.2532303, similarity=11.80944]
RECALL MISS: 72311 1996_Summer_Olympics]

4: 72311 1996_Summer_Olympics

```
[winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.22617276, final_local_scores=-0.29926634,
pem=-3.2532303, similarity=12.850581]
    RECALL MISS: 72311 1996_Summer_Olympics]
1: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.4578191,
final_local_scores=1.5459425, pem=0.0, similarity=50.22372]
2: 46827 Jesse_Owens
    [pem_pos=0, final_scores=1.1802998,
final_local_scores=1.233433, pem=0.0, similarity=40.89781]
3: 3295 Bonn
    [pem_pos=0, final_scores=0.3937651,
final_local_scores=0.45153773, pem=0.0, similarity=17.564499]
4: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.4820302,
final_local_scores=0.5116146, pem=-0.61114657,
similarity=22.680824]
5: 27566 Summer_Olympic_Games
    [pem_pos=2, final_scores=0.36454314,
final_local_scores=0.35074037, pem=-3.2532303,
similarity=32.248077]
6: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.1013185,
final_local_scores=1.1515992, pem=0.0, similarity=38.455727]
7: 46827 Jesse_Owens
    [pem_pos=0, final_scores=0.68268913,
final_local_scores=0.6854547, pem=0.0, similarity=24.54505]
8: 3354 Berlin
    [pem_pos=0, final_scores=0.50203425,
final_local_scores=0.5294899, pem=-0.15483057,
similarity=20.73274]
9: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.6687629,
final_local_scores=1.7763667, pem=0.0, similarity=57.10004]
10: 11867 Germany
    [pem_pos=1, final_scores=0.72183,
final_local_scores=0.78100485, pem=-1.0465823,
similarity=33.087933]
11: 3354 Berlin
    [pem_pos=0, final_scores=0.6554248,
final_local_scores=0.6970457, pem=-0.15483057,
similarity=25.73294]
12: 27566 Summer_Olympic_Games
    [pem_pos=2, final_scores=0.32875854,
final_local_scores=0.31165105, pem=-3.2532303,
similarity=31.081573]
13: 46827 Jesse_Owens
    [pem_pos=0, final_scores=1.3114395,
final_local_scores=1.3766832, pem=0.0, similarity=45.172688]
14: 5042916 Canada
    [pem_pos=0, final_scores=0.788057,
final_local_scores=0.85665894, pem=-0.11662793,
```

```
similarity=30.288366]
15: 67116 Donovan_Bailey
    [pem_pos=0, final_scores=0.55292267,
final_local_scores=0.60651654, pem=0.0, similarity=22.189377]
16: 71410 Frankie_Fredericks
    [pem_pos=0, final_scores=0.8424005,
final_local_scores=0.9423247, pem=0.0, similarity=32.21056]
17: 72311 1996_Summer_Olympics
    [pem_pos=14, final_scores=0.341515,
final_local_scores=0.33264714, pem=-4.337497,
similarity=37.604553]
18: 653621 Michael_Johnson_(sprinter)
    [pem_pos=0, final_scores=0.55547476,
final_local_scores=0.57219344, pem=-0.36135122,
similarity=23.13019]
19: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.6122956,
final_local_scores=1.7146847, pem=0.0, similarity=55.259327]
20: 27566 Summer_Olympic_Games
    [pem_pos=2, final_scores=0.39896232,
final_local_scores=0.38833803, pem=-3.2532303,
similarity=33.370068]
21: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5117907,
final_local_scores=0.5441235, pem=-0.61114657,
similarity=23.650953]
22: 67299 Allan_Wells
    [pem_pos=0, final_scores=0.6120263,
final_local_scores=0.66123813, pem=0.0, similarity=23.822378]
23: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.6023418,
final_local_scores=1.7038116, pem=0.0, similarity=54.934853]
24: 31717 United_Kingdom
    [pem_pos=1, final_scores=0.45984685,
final_local_scores=0.4873827, pem=-0.9694005,
similarity=23.905935]
25: 46827 Jesse_Owens
    [pem_pos=0, final_scores=1.2814466,
final_local_scores=1.3439205, pem=0.0, similarity=44.19498]
26: 5334607 Africa
    [pem_pos=0, final_scores=0.5368181,
final_local_scores=0.6282225, pem=-0.13197383,
similarity=23.554821]
27: 3434750 United_States
    [pem_pos=0, final_scores=0.5192335,
final_local_scores=0.5546849, pem=-0.073457584,
similarity=21.04209]
28: 9239 Europe
    [pem_pos=0, final_scores=0.6228992,
final_local_scores=0.699802, pem=-0.086338006,
similarity=25.44272]
29: 46827 Jesse_Owens
```

```
[pem_pos=0, final_scores=1.3089719,
final_local_scores=1.3739877, pem=0.0, similarity=45.092243]
30: 67127 Carl_Lewis
    [pem_pos=0, final_scores=0.54091126,
final_local_scores=0.5535059, pem=0.0, similarity=20.607433]
31: 3708 Brussels
    [pem_pos=0, final_scores=0.63404334,
final_local_scores=0.7043591, pem=-0.025440974,
similarity=25.247547]
```

Sample: b'685_CRICKET'

CRICKET - NEW CAPTAIN [1 1 TENDULKAR 1] UPSTAGED BY 120 FROM [2 2 JAYASURIYA 2] . [3 3 COLOMBO 3] 1996-08-28 [4 4 Sachin_Tendulkar 4] marked his debut as [5 5 Indian 5] captain with a patient 110 on Wednesday , but was upstaged by dashing [6 6 Sri_Lankan 6] opener [7 7 Sanath_Jayasuriya 7] whose 120 steered the world champions to a nine-wicket Singer Cup win . [8 8 Sri_Lanka 8] , playing in front of their home crowd for the first time since winning the [9 9 World_Cup 9] last March , comfortably passed [10 10 India 10] 's modest 226-5 from 50 overs in 44.2 overs . The devastating opening pair of [11 11 Jayasuriya 11] and [12 12 Romesh_Kaluwitharana 12] shared a fine first wicket stand of 129 to the delight of the 25,000 fans . [13 13 Jayasuriya 13] , whose first 50 included three sixes and three fours , went on to an unbeaten 120 and the man-of-the-match award . [14 14 Kaluwitharana 14] , slow in comparison , was bowled by [15 15 Tendulkar 15] for 53 , but [16 16 Aravinda_de_Silva 16] with 49 not out helped see [17 17 Sri_Lanka 17] home . Earlier , [18 18 Tendulkar 18] completed his ninth century in one-day cricket , taking 138 balls to do it before being run out . The rest of the [19 19 Indian 19] batting was generally tied down by brilliant fielding and some fairly tight bowling , although ex-captain Mohamed_Azharuddin chipped in with 58 , adding 129 with [20 20 Tendulkar 20] off 28 overs , before being stumped . The next match in the four-nation tournament is on Friday when [21 21 Sri_Lanka 21] play [22 22 Australia 22] in a repeat of the [23 23 World_Cup 23] final in [24 24 Lahore 24] where [25 25 Sri_Lanka 25] won by seven wickets .

```
1: 57570 Sachin_Tendulkar
2: 397266 Sanath_Jayasuriya
3: 56636 Colombo
4: 57570 Sachin_Tendulkar
5: 14533 India
6: 397246 Sri_Lanka_national_cricket_team
7: 397266 Sanath_Jayasuriya
8: 397246 Sri_Lanka_national_cricket_team
9: 7239 Cricket_World_Cup
10: 407754 India_national_cricket_team
11: 397266 Sanath_Jayasuriya
12: 894385 Romesh_Kaluwitharana
13: 397266 Sanath_Jayasuriya
```

```
14: 894385 Romesh_Kaluwitharana
15: 57570 Sachin_Tendulkar
16: 30873370 Aravinda_de_Silva
17: 397246 Sri_Lanka_national_cricket_team
18: 57570 Sachin_Tendulkar
19: 407754 India_national_cricket_team
20: 57570 Sachin_Tendulkar
21: 397246 Sri_Lanka_national_cricket_team
22: 374833 Australia_national_cricket_team
23: 7239 Cricket_World_Cup
24: 125315 Lahore
25: 397246 Sri_Lanka_national_cricket_team
1: 57570 Sachin_Tendulkar
    [pem_pos=0, final_scores=0.9681657,
final_local_scores=0.97524625, pem=0.0, similarity=33.193005]
2: 397266 Sanath_Jayasuriya
    [pem_pos=0, final_scores=0.46116596,
final_local_scores=0.4263813, pem=0.0, similarity=16.813782]
3: 56636 Colombo
    [pem_pos=0, final_scores=0.58167404,
final_local_scores=0.62671906, pem=-0.19393882,
similarity=23.84693]
4: 57570 Sachin_Tendulkar
    [pem_pos=0, final_scores=0.6920254,
final_local_scores=0.67360497, pem=0.0, similarity=24.19143]
5: 14533 India
    [pem_pos=0, final_scores=0.48429918,
final_local_scores=0.5741896, pem=-0.618294,
similarity=24.587055]
6: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.65414035,
final_local_scores=0.6179132, pem=-1.9963025,
similarity=33.385677]
7: 397266 Sanath_Jayasuriya
    [pem_pos=0, final_scores=0.86370116,
final_local_scores=0.8660899, pem=0.0, similarity=29.935562]
8: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.7691112,
final_local_scores=0.74350137, pem=-2.2667584,
similarity=38.60426]
9: 7239 Cricket_World_Cup
    [pem_pos=4, final_scores=0.544077,
final_local_scores=0.52674174, pem=-3.1731894,
similarity=37.065033]
10: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.79232126,
final_local_scores=0.7629223, pem=-3.5152702,
similarity=45.973423]
11: 397266 Sanath_Jayasuriya
    [pem_pos=0, final_scores=0.56352985,
final_local_scores=0.5381983, pem=0.0, similarity=20.150623]
12: 894385 Romesh_Kaluwitharana
```

```
[pem_pos=0, final_scores=0.7057082,
final_local_scores=0.76554424, pem=0.0, similarity=26.93508]
13: 397266 Sanath_Jayasuriya
    [pem_pos=0, final_scores=0.6504968,
final_local_scores=0.6331965, pem=0.0, similarity=22.98556]
14: 894385 Romesh_Kaluwitharana
    [pem_pos=0, final_scores=0.87872976,
final_local_scores=0.954544, pem=0.0, similarity=32.57521]
15: 57570 Sachin_Tendulkar
    [pem_pos=0, final_scores=0.64592135,
final_local_scores=0.6232433, pem=0.0, similarity=22.688536]
16: 30873370 Aravinda_de_Silva
    [pem_pos=0, final_scores=0.58846265,
final_local_scores=0.5915868, pem=0.0, similarity=21.743843]
17: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.7468251,
final_local_scores=0.71915716, pem=-2.2667584,
similarity=37.87778]
18: 57570 Sachin_Tendulkar
    [pem_pos=0, final_scores=0.63888466,
final_local_scores=0.6155568, pem=0.0, similarity=22.459154]
19: 407754 India_national_cricket_team
    [pem_pos=10, final_scores=0.5690946,
final_local_scores=0.5190811, pem=-3.95575,
similarity=41.092113]
20: 57570 Sachin_Tendulkar
    [pem_pos=0, final_scores=0.55531377,
final_local_scores=0.5242682, pem=0.0, similarity=19.734922]
21: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.72155505,
final_local_scores=0.69155353, pem=-2.2667584,
similarity=37.054035]
22: 374833 Australia_national_cricket_team
    [pem_pos=1, final_scores=0.8464909,
final_local_scores=0.834062, pem=-3.4989398,
similarity=48.007565]
23: 7239 Cricket_World_Cup
    [pem_pos=4, final_scores=0.56958383,
final_local_scores=0.55460405, pem=-3.1731894,
similarity=37.8965]
24: 125315 Lahore
    [pem_pos=0, final_scores=0.43631017,
final_local_scores=0.51640564, pem=0.0, similarity=19.500288]
25: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.6683918,
final_local_scores=0.6334807, pem=-2.2667584,
similarity=35.321026]
```

Sample: b'686_CRICKET'

CRICKET - [1 1 SRI LANKA 1] BEAT [2 2 INDIA 2] BY 9 WICKETS
IN ONE-DAY MATCH . [3 3 COLOMBO 3] 1996-08-28 [4 4 Sri Lanka 4] beat [5 5 India 5] by nine wickets in the second match of

the [6 6 Singer World Series 6] one-day (50 overs) cricket tournament on Monday . Scores : [7 7 India 7] 226-5 in 50 overs , [8 8 Sri Lanka 8] 230-1 in 44.2 overs .

```

1: 397246 Sri_Lanka_national_cricket_team
2: 407754 India_national_cricket_team
3: 56636 Colombo
4: 397246 Sri_Lanka_national_cricket_team
5: 407754 India_national_cricket_team
6: 28357832 Singer_World_Series
7: 407754 India_national_cricket_team
8: 397246 Sri_Lanka_national_cricket_team
1: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.6683561,
final_local_scores=0.6175973, pem=-2.2667584,
similarity=34.84703]
2: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.5862557,
final_local_scores=0.52956915, pem=-3.5152702,
similarity=39.0097]
3: 56636 Colombo
    [pem_pos=0, final_scores=0.5230028,
final_local_scores=0.57671195, pem=-0.19393882,
similarity=22.354618]
4: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.76089936,
final_local_scores=0.71868664, pem=-2.2667584,
similarity=37.86374]
5: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.7258422,
final_local_scores=0.6820462, pem=-3.5152702,
similarity=43.559917]
6: 28357832 Singer_World_Series
    [pem_pos=0, final_scores=0.63473314,
final_local_scores=0.6992774, pem=0.0, similarity=24.957546]
7: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.7448159,
final_local_scores=0.70277214, pem=-3.5152702,
similarity=44.17842]
8: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.8430472,
final_local_scores=0.8084207, pem=-2.2667584,
similarity=40.541584]
```

Sample: b'687_CRICKET'

CRICKET - [1 1 INDIA 1] V [2 2 SRI LANKA 2] SCOREBOARD . [3 3 COLOMBO 3] 1996-08-28 Scoreboard of the second [4 4 Singer Wo rld Series 4] cricket match between [5 5 India 5] and [6 6 Sri Lanka 6] on Wednesday : [7 7 India 7] A. Jadeja run out 0 S. Tendulkar run out 110 S. Ganguly c de Silva b [8 8 Dharmasena 8] 16 M. Azharuddin st [9 9 Kaluwitharana 9] b [10 10 Jayasuriya 10] 58 V. Kamblu run out 18 [11 11 R. Dravid 11] not out 7 J. Srinath not out 1 Extras (b-1 lb-3 w-9 nb-3) 16 Total

(5 wickets , 50 overs) 226 Fall of wickets : 1-4 2-57 3-186
4-217 5-217 . Did not bat : [12 12 A. Kumble 12] , N.
Mongia , V. Prasad , A. Kapoor . Bowling : Vass 9-2-35-0 , [13
13 Pushpakumara 13] 6-0-23-0 , [14 14 Dharmasena 14]
10-0-59-1 [15 15 Muralitharan 15] 10-0-42-0 , [16 16 Jayasuri
ya 16] 10-1-39-1 , de Silva 5-0-24-0 . [17 17 Sri Lanka 17]
S. Jayasuriya not out 120 R. Kaluwitharana b [18 18 Tendulkar
18] 53 A.de Silva not out 49 Extras (lb-3 nb-3 w-2) 8 Total
(for one wicket - 44.2 overs) 230 Fall of wicket : 1-129
Did not bat : [19 19 Arjuna Ranatunga 19] , [20 20 Asanka Gur
usinha 20] , Hashan Tillakaratne , [21 21 Roshan Mahanama 21]
, Kumara Dharmasena , [22 22 Chaminda Vaas 22] , [23 23 Muthi
ah Muralitharan 23] , [24 24 Ravindra Pushpakumara 24]
Bowling : [25 25 Kumble 25] 10-1-40-0 , [26 26 Prasad 26]
6-0-47-0 , [27 27 Srinath 27] 8-0-33-0 , [28 28 Tendulkar 28]
6-0-29-1 , [29 29 Kapoor 29] 10-2-51-0 , [30 30 Jadeja 30]
2.2-0-13-0 , [31 31 Ganguly 31] 2-0-14-0 Result : [32 32 Sri L
anka 32] won by 9 wickets Man-of-the-Match : [33 33 Sanath Ja
yasuriya 33]

1: 407754 India_national_cricket_team
2: 397246 Sri_Lanka_national_cricket_team
3: 56636 Colombo
4: 28357832 Singer_World_Series
5: 407754 India_national_cricket_team
6: 397246 Sri_Lanka_national_cricket_team
7: 407754 India_national_cricket_team
8: 3105984 Kumar_Dharmasena
9: 894385 Romesh_Kaluwitharana
10: 397266 Sanath_Jayasuriya
11: 252735 Rahul_Dravid
12: 274600 Anil_Kumble
13: 3140804 Ravindra_Pushpakumara
14: 3105984 Kumar_Dharmasena
15: 20905 Muttiah_Muralitharan
16: 397266 Sanath_Jayasuriya
17: 397246 Sri_Lanka_national_cricket_team
18: 57570 Sachin_Tendulkar
19: 1160705 Arjuna_Ranatunga
20: 6699670 Asanka_Gurusinha
21: 1671083 Roshan_Mahanama
22: 198062 Chaminda_Vaas
23: 20905 Muttiah_Muralitharan
24: 3140804 Ravindra_Pushpakumara
25: 274600 Anil_Kumble
26: 3060336 Venkatesh_Prasad
27: 492474 Javagal_Srinath
28: 57570 Sachin_Tendulkar
29: 3965321 Aashish_Kapoor
30: 3369198 Ajay_Jadeja
31: 529681 Sourav_Ganguly
32: 397246 Sri_Lanka_national_cricket_team
33: 397266 Sanath_Jayasuriya

```
1: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.6791794,
final_local_scores=0.64634895, pem=-3.5152702,
similarity=42.49464]
2: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.6401349,
final_local_scores=0.61114997, pem=-2.2667584,
similarity=34.654633]
3: 56636 Colombo
    [pem_pos=0, final_scores=0.5362086,
final_local_scores=0.5964971, pem=-0.19393882,
similarity=22.945047]
4: 28357832 Singer_World_Series
    [pem_pos=0, final_scores=0.65450174,
final_local_scores=0.70517385, pem=0.0, similarity=25.133507]
5: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.74064595,
final_local_scores=0.7134919, pem=-3.5152702,
similarity=44.49832]
6: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.79297334,
final_local_scores=0.7781028, pem=-2.2667584,
similarity=39.636837]
7: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.7772588,
final_local_scores=0.7534859, pem=-3.5152702,
similarity=45.691822]
8: 3105984 Kumar_Dharmasena
    [pem_pos=0, final_scores=0.5123812,
final_local_scores=0.54704976, pem=0.0, similarity=20.414768]
9: 894385 Romesh_Kaluwitharana
    [pem_pos=0, final_scores=0.5768253,
final_local_scores=0.6179689, pem=0.0, similarity=22.531137]
10: 397266 Sanath_Jayasuriya
    [pem_pos=0, final_scores=0.4297129,
final_local_scores=0.38898236, pem=0.0, similarity=15.697722]
11: 252735 Rahul_Dravid
    [pem_pos=0, final_scores=0.54524165,
final_local_scores=0.52441746, pem=0.0, similarity=19.739376]
12: 274600 Anil_Kumble
    [pem_pos=0, final_scores=0.7123773,
final_local_scores=0.6998373, pem=0.0, similarity=24.974255]
13: 3140804 Ravindra_Pushpakumara
    [pem_pos=0, final_scores=0.3505085,
final_local_scores=0.4191414, pem=0.0, similarity=16.597729]
14: 3105984 Kumar_Dharmasena
    [pem_pos=0, final_scores=0.5194844,
final_local_scores=0.5548089, pem=0.0, similarity=20.646318]
15: 20905 Muttiah_Muralitharan
    [pem_pos=0, final_scores=0.7381987,
final_local_scores=0.7226778, pem=0.0, similarity=25.65586]
16: 397266 Sanath_Jayasuriya
```

```
[pem_pos=0, final_scores=0.533162,
final_local_scores=0.5019848, pem=0.0, similarity=19.06994]
17: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.6765707,
final_local_scores=0.6509506, pem=-2.2667584,
similarity=35.84236]
18: 57570 Sachin_Tendulkar
    [pem_pos=0, final_scores=0.47822136,
final_local_scores=0.4448743, pem=0.0, similarity=17.365648]
19: 1160705 Arjuna_Ranatunga
    [pem_pos=0, final_scores=0.6502849,
final_local_scores=0.6833679, pem=0.0, similarity=24.482775]
20: 6699670 Asanka_Gurusinha
    [pem_pos=0, final_scores=0.64643997,
final_local_scores=0.71437395, pem=0.0, similarity=25.408056]
21: 1671083 Roshan_Mahanama
    [pem_pos=0, final_scores=0.70746577,
final_local_scores=0.7650665, pem=0.0, similarity=26.920824]
22: 198062 Chaminda_Vaas
    [pem_pos=0, final_scores=0.759564,
final_local_scores=0.765449, pem=0.0, similarity=26.932238]
23: 20905 Muttiah_Muralitharan
    [pem_pos=0, final_scores=0.713824,
final_local_scores=0.69605213, pem=0.0, similarity=24.861298]
24: 3140804 Ravindra_Pushpakumara
    [pem_pos=0, final_scores=0.40927643,
final_local_scores=0.48333645, pem=0.0, similarity=18.513437]
25: 274600 Anil_Kumble
    [pem_pos=0, final_scores=0.8057504,
final_local_scores=0.8018332, pem=0.0, similarity=28.018015]
26: 3060336 Venkatesh_Prasad
    [pem_pos=1, final_scores=0.6534912,
final_local_scores=0.6712831, pem=-2.55979,
similarity=38.042675]
27: 492474 Javagal_Srinath
    [pem_pos=2, final_scores=0.52578896,
final_local_scores=0.51873165, pem=-1.1409599,
similarity=25.774418]
28: 57570 Sachin_Tendulkar
    [pem_pos=0, final_scores=0.52279,
final_local_scores=0.49355882, pem=0.0, similarity=18.818493]
29: 3965321 Aashish_Kapoor
    [pem_pos=2, final_scores=0.52402455,
final_local_scores=0.62598217, pem=-3.2394922,
similarity=40.38713]
30: 3369198 Ajay_Jadeja
    [pem_pos=1, final_scores=0.41821843,
final_local_scores=0.44980782, pem=-1.5870132,
similarity=26.1433]
31: 529681 Sourav_Ganguly
    [pem_pos=0, final_scores=0.4995275,
final_local_scores=0.47984332, pem=0.0, similarity=18.409195]
```

```
32: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.75667495,
final_local_scores=0.73845226, pem=-2.2667584,
similarity=38.453587]
33: 397266 Sanath_Jayasuriya
    [pem_pos=0, final_scores=0.6603485,
final_local_scores=0.6409167, pem=0.0, similarity=23.215946]
```

Sample: b'688_CRICKET'

CRICKET - [1 1 INDIA 1] WIN TOSS AND BAT AGAINST [2 2 SRI LA NKA 2] . [3 3 COLOMBO 3] 1996-08-28 [4 4 India 4] won the toss and elected to bat against [5 5 Sri Lanka 5] in the second day-night limited overs cricket match of the [6 6 Singer World Series 6] tournament on Wednesday . Teams : [7 7 India 7] - [8 8 Sachin Tendulkar 8] (captain) , [9 9 Anil Kumble 9] , [10 10 Ajay Jadeja 10] , [11 11 Sourav Ganguly 11] , Mohamed Azharuddin , [12 12 Vinod Kambli 12] , [13 13 Rahul Dravid 13] , [14 14 Nayan Mongia 14] , [15 15 Javagal Srinath 15] , [16 16 Venkatesh Prasad 16] , Ashish Kapoor . [17 17 Sri Lanka 17] - [18 18 Arjuna Ranatunga 18] (captain) , [19 19 Sanath Jayasuriya 19] , [20 20 Romesh Kaluwitharana 20] , [21 21 Asanka Gurusinha 21] , [22 22 Aravinda de Silva 22] , Hashan Tillekeratne , [23 23 Roshan Mahanama 23] , Kumara Dharmasena , [24 24 Chaminda Vaas 24] , [25 25 Muttiah Muralitharan 25] , [26 26 Ravindra Pushpakumara 26] .

```
1: 407754 India_national_cricket_team
2: 397246 Sri_Lanka_national_cricket_team
3: 56636 Colombo
4: 407754 India_national_cricket_team
5: 397246 Sri_Lanka_national_cricket_team
6: 28357832 Singer_World_Series
7: 407754 India_national_cricket_team
8: 57570 Sachin_Tendulkar
9: 274600 Anil_Kumble
10: 3369198 Ajay_Jadeja
11: 529681 Sourav_Ganguly
12: 564046 Vinod_Kambli
13: 252735 Rahul_Dravid
14: 3059959 Nayan_Mongia
15: 492474 Javagal_Srinath
16: 3060336 Venkatesh_Prasad
17: 397246 Sri_Lanka_national_cricket_team
18: 1160705 Arjuna_Ranatunga
19: 397266 Sanath_Jayasuriya
20: 894385 Romesh_Kaluwitharana
21: 6699670 Asanka_Gurusinha
22: 30873370 Aravinda_de_Silva
23: 1671083 Roshan_Mahanama
24: 198062 Chaminda_Vaas
25: 20905 Muttiah_Muralitharan
26: 3140804 Ravindra_Pushpakumara
1: 407754 India_national_cricket_team
```

```
[pem_pos=1, final_scores=0.6702153,
final_local_scores=0.63439894, pem=-3.5152702,
similarity=42.13803]
2: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.42837507,
final_local_scores=0.3804357, pem=-2.2667584,
similarity=27.769657]
3: 56636 Colombo
    [pem_pos=0, final_scores=0.5155476,
final_local_scores=0.5775487, pem=-0.19393882,
similarity=22.379587]
4: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.667288,
final_local_scores=0.63120127, pem=-3.5152702,
similarity=42.042606]
5: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.82119113,
final_local_scores=0.8095276, pem=-2.2667584,
similarity=40.574615]
6: 28357832 Singer_World_Series
    [pem_pos=0, final_scores=0.7256317,
final_local_scores=0.7841364, pem=0.0, similarity=27.489908]
7: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.73244554,
final_local_scores=0.702376, pem=-3.5152702,
similarity=44.1666]
8: 57570 Sachin_Tendulkar
    [pem_pos=0, final_scores=0.6673451,
final_local_scores=0.65054435, pem=0.0, similarity=23.503254]
9: 274600 Anil_Kumble
    [pem_pos=0, final_scores=0.7887028,
final_local_scores=0.7837401, pem=0.0, similarity=27.47808]
10: 3369198 Ajay_Jadeja
    [pem_pos=0, final_scores=0.5404634,
final_local_scores=0.58283335, pem=0.0, similarity=21.482622]
11: 529681 Sourav_Ganguly
    [pem_pos=0, final_scores=0.57182604,
final_local_scores=0.55498046, pem=0.0, similarity=20.651436]
12: 564046 Vinod_Kambli
    [pem_pos=0, final_scores=0.6382812,
final_local_scores=0.6913086, pem=0.0, similarity=24.719742]
13: 252735 Rahul_Dravid
    [pem_pos=0, final_scores=0.71098095,
final_local_scores=0.7024212, pem=0.0, similarity=25.051363]
14: 3059959 Nayan_Mongia
    [pem_pos=0, final_scores=0.5708231,
final_local_scores=0.61733484, pem=0.0, similarity=22.512217]
15: 492474 Javagal_Srinath
    [pem_pos=0, final_scores=0.8163392,
final_local_scores=0.83410543, pem=0.0, similarity=28.981083]
16: 3060336 Venkatesh_Prasad
    [pem_pos=0, final_scores=0.8400715,
```

```
final_local_scores=0.87680775, pem=0.0, similarity=30.255405]
17: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.7572201,
final_local_scores=0.7396489, pem=-2.2667584,
similarity=38.489296]
18: 1160705 Arjuna_Ranatunga
    [pem_pos=0, final_scores=0.6593855,
final_local_scores=0.6932754, pem=0.0, similarity=24.778433]
19: 397266 Sanath_Jayasuriya
    [pem_pos=0, final_scores=0.84729797,
final_local_scores=0.85095155, pem=0.0, similarity=29.483805]
20: 894385 Romesh_Kaluwitharana
    [pem_pos=0, final_scores=0.6987229,
final_local_scores=0.75349414, pem=0.0, similarity=26.575481]
21: 6699670 Asanka_Gurusinha
    [pem_pos=0, final_scores=0.6037303,
final_local_scores=0.6694675, pem=0.0, similarity=24.067959]
22: 30873370 Aravinda_de_Silva
    [pem_pos=0, final_scores=0.71138024,
final_local_scores=0.7158997, pem=0.0, similarity=25.453588]
23: 1671083 Roshan_Mahanama
    [pem_pos=0, final_scores=0.6949739,
final_local_scores=0.75179166, pem=0.0, similarity=26.524675]
24: 198062 Chaminda_Vaas
    [pem_pos=0, final_scores=0.8001978,
final_local_scores=0.811681, pem=0.0, similarity=28.311892]
25: 20905 Muttiah_Muralitharan
    [pem_pos=0, final_scores=0.8117451,
final_local_scores=0.8072644, pem=0.0, similarity=28.180092]
26: 3140804 Ravindra_Pushpakumara
    [pem_pos=0, final_scores=0.5072544,
final_local_scores=0.60105836, pem=0.0, similarity=22.026493]
```

Sample: b'689_PRESS'

PRESS DIGEST - [1 1 ANGOLA 1] - AUG 28 . [2 2 LUANDA 2]
1996-08-28 These are the leading stories in the [3 3 Angolan 3]
] press on Wednesday . [4 4 Reuters 4] has not verified these
stories and does not vouch for their accuracy . JORNAL DE
ANGOLA - [5 5 Princeton Lyman 5] , the [6 6 U.S. 6] Under-
Secretary of State for International Organisations , will on
Wednesday continue his work in [7 7 Angola 7] visiting [8 8 Ba-
ilundo 8] , where he should be received by [9 9 Jonas Savimbi
9] , leader of [10 10 Unita 10] . On Tuesday [11 11 Lyman 11]
participated in a meeting of a joint-commission where he
considered that the [12 12 Angolan 12] politicians should
advance faster and find a way to cooperate . In his opinion
the quartering of [13 13 Unita 13] forces must be concluded
in all the [14 14 Angolan 14] territory and the troops must
be selected and integrated in the armed forces , the
government forces must be concentrated in the principal units
and the free circulation of people and goods must be reality
in all the country .

```
1: 701 Angola
2: 18402 Luanda
3: 701 Angola
4: 18998750 Reuters
5: 18298102 Princeton_Lyman
6: 3434750 United_States
7: 701 Angola
8: 13699261 Bailundo
9: 264097 Jonas_Savimbi
10: 245423 UNITA
11: 18298102 Princeton_Lyman
12: 701 Angola
13: 245423 UNITA
14: 701 Angola
1: 701 Angola
    [pem_pos=0, final_scores=0.52354354,
final_local_scores=0.48507386, pem=-0.39249876,
similarity=20.699755]
2: 18402 Luanda
    [pem_pos=0, final_scores=0.4087147,
final_local_scores=0.400774, pem=0.0, similarity=16.049608]
3: 701 Angola
    [pem_pos=0, final_scores=0.54725355,
final_local_scores=0.5109734, pem=-0.020324398,
similarity=19.448706]
4: 18998750 Reuters
    [pem_pos=0, final_scores=0.27581578,
final_local_scores=0.3455239, pem=0.0, similarity=14.400835]
5: 18298102 Princeton_Lyman
    [pem_pos=0, final_scores=0.4139437,
final_local_scores=0.52804667, pem=0.0, similarity=19.84768]
6: 3434750 United_States
    [pem_pos=0, final_scores=0.4869737,
final_local_scores=0.5223929, pem=-0.14075686,
similarity=20.444416]
7: 701 Angola
    [pem_pos=0, final_scores=0.5830497,
final_local_scores=0.5500753, pem=-0.39249876,
similarity=22.639526]
8: 13699261 Bailundo
    [pem_pos=0, final_scores=0.41796052,
final_local_scores=0.518691, pem=0.0, similarity=19.568487]
9: 264097 Jonas_Savimbi
    [pem_pos=0, final_scores=0.6914987,
final_local_scores=0.73761004, pem=0.0, similarity=26.101467]
10: 245423 UNITA
    [pem_pos=0, final_scores=0.45458466,
final_local_scores=0.45674193, pem=0.0, similarity=17.719803]
11: 18298102 Princeton_Lyman
    [pem_pos=0, final_scores=0.37487984,
final_local_scores=0.48537534, pem=0.0, similarity=18.574282]
12: 701 Angola
```

```
[pem_pos=0, final_scores=0.5058311,
final_local_scores=0.46572572, pem=-0.020324398,
similarity=18.098425]
13: 245423 UNITA
[pem_pos=0, final_scores=0.57760924,
final_local_scores=0.59112763, pem=0.0, similarity=21.73014]
14: 701 Angola
[pem_pos=0, final_scores=0.4997059,
final_local_scores=0.45903486, pem=-0.020324398,
similarity=17.898756]
```

Sample: b'690_S.AFRICAN'

S.AFRICAN TRUTH BODY TO SUMMON APARTHEID POLICE . [1 1 CAPE T OWN 1] 1996-08-28 [2 2 South Africa 2] 's Truth and Reconciliation Commission said on Wednesday it would subpoena persons accused of human rights violations to appear before it . " We can subpoena anyone we want to , even the president of the country , " spokesman John Allen told [3 3 Reuters 3] . " Subpoenas are due to be served on a number of people this week . " Media reports have speculated that the commission , which is trying to heal the wounds of apartheid by confronting the past , could subpoena apartheid-era President [4 4 P.W. Botha 4] and former police generals Basie Smit and Johan Van Der Merwe . In submissions last week to the commission [5 5 National Party 5] leader and former president F.W. De Klerk said he had received no co-operation from [6 6 Botha 6] in compiling his party 's report . Since it began work in April the commission has been hearing harrowing tales from the victims of apartheid-era abuses , by both the white minority regime and its opponents . It also wants to hear from those who committed the abuses , to whom it can offer amnesty in return for frankness . Hopes that reformed perpetrators would come forward voluntarily have faded but the commission has the legal power to force them to appear . Allen declined to give say who would be subpoenaed . " At the moment we have a preliminary list of less than 10 people , but this is just the beginning , " he said . The commission was set up last year to probe 30 years of human-rights violations during the apartheid era . It is chaired by [7 7 Nobel Peace 7] winner , retired Archbishop [8 8 Desmond Tutu 8] . Allen said the commission could announce the names of subpoenaed persons on Monday next week .

```
1: 6653 Cape_Town
2: 17416221 South_Africa
3: 18998750 Reuters
4: 383252 P._W._Botha
5: 390975 National_Party_(South_Africa)
6: 383252 P._W._Botha
7: 26230922 Nobel_Peace_Prize
8: 8401973 Desmond_Tutu
1: 6653 Cape_Town
[pem_pos=0, final_scores=0.45804006,
```

```
final_local_scores=0.48562855, pem=0.0, similarity=18.581837]
2: 17416221 South_Africa
    [pem_pos=0, final_scores=0.45332414,
final_local_scores=0.489406, pem=-0.26251382,
similarity=20.122154]
3: 18998750 Reuters
    [pem_pos=0, final_scores=0.32280183,
final_local_scores=0.3728311, pem=0.0, similarity=15.215736]
4: 383252 P._W._Botha
    [pem_pos=0, final_scores=0.741811,
final_local_scores=0.74342936, pem=0.0, similarity=26.275127]
5: 390975 National_Party_(South_Africa)
    [pem_pos=0, final_scores=0.4665215,
final_local_scores=0.46384943, pem=-0.99382544,
similarity=23.336483]
6: 383252 P._W._Botha
    [pem_pos=0, final_scores=0.58160675,
final_local_scores=0.5903589, pem=0.0, similarity=21.707201]
7: 26230922 Nobel_Peace_Prize
    [pem_pos=0, final_scores=0.3066998,
final_local_scores=0.34473655, pem=0.0, similarity=14.377339]
8: 8401973 Desmond_Tutu
    [pem_pos=0, final_scores=0.51488477,
final_local_scores=0.5579791, pem=0.0, similarity=20.740923]
```

Sample: b'691_TURKISH'

TURKISH AIRPLANE LANDS IN [1 1 SOFIA 1] ON BOMB THREAT . [2 2 SOFIA 2] 1996-08-28 A Turkish airliner on flight from [3 3 Is tanbul 3] to [4 4 Vienna 4] on Wednesday landed in emergency at [5 5 Sofia 5] airport after receiving a bomb threat , said an airport official . " The plane landed at [6 6 Sofia 6] airport at 1503 (1203 [7 7 GMT 7]) after receiving a signal that there is an explosive on board , " the official , who declined to be named told [8 8 Reuters 8] . The plane , surrounded by 11 fire-engines , is being checked for explosives at the moment . Nothing has been found so far , added the official . In March a [9 9 Turkish_Cypriot 9] airliner hijacked while on a flight from northern [10 10 Cypr us 10] to [11 11 Istanbul 11] landed in [12 12 Sofia 12] airport to refuel before landing in [13 13 Munich 13] , where the hijacker was arrested .

```
1: 57644 Sofia
2: 57644 Sofia
3: 3391396 Istanbul
4: 55866 Vienna
5: 57644 Sofia
6: 57644 Sofia
7: 12701 Greenwich_Mean_Time
8: 18998750 Reuters
9: 934562 Turkish_Cypriots
10: 5593 Cyprus
11: 3391396 Istanbul
```

```
12: 57644 Sofia
13: 19058 Munich
1: 57644 Sofia
    [pem_pos=0, final_scores=0.45486593,
final_local_scores=0.39343292, pem=-0.10680699,
similarity=16.41137]
2: 57644 Sofia
    [pem_pos=0, final_scores=0.5973193,
final_local_scores=0.54904157, pem=-0.10680699,
similarity=21.05504]
3: 3391396 Istanbul
    [pem_pos=0, final_scores=0.47430748,
final_local_scores=0.42987126, pem=-0.025284406,
similarity=17.05543]
4: 55866 Vienna
    [pem_pos=0, final_scores=0.5973864,
final_local_scores=0.58224064, pem=-0.12517071,
similarity=22.145632]
5: 57644 Sofia
    [pem_pos=0, final_scores=0.58254075,
final_local_scores=0.53289825, pem=-0.10680699,
similarity=20.573292]
6: 57644 Sofia
    [pem_pos=0, final_scores=0.5425265,
final_local_scores=0.48918873, pem=-0.10680699,
similarity=19.268913]
7: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.6409476,
final_local_scores=0.75616676, pem=0.0, similarity=26.655237]
8: 18998750 Reuters
    [pem_pos=0, final_scores=0.36712503,
final_local_scores=0.4299628, pem=0.0, similarity=16.92066]
9: 934562 Turkish_Cypriots
    [pem_pos=0, final_scores=0.26934344,
final_local_scores=0.27824295, pem=0.0, similarity=12.393038]
10: 5593 Cyprus
    [pem_pos=0, final_scores=0.5385451,
final_local_scores=0.5453934, pem=-0.3752463,
similarity=22.405987]
11: 3391396 Istanbul
    [pem_pos=0, final_scores=0.42186284,
final_local_scores=0.37258345, pem=-0.025284406,
similarity=15.345846]
12: 57644 Sofia
    [pem_pos=0, final_scores=0.61643994,
final_local_scores=0.569928, pem=-0.10680699,
similarity=21.678333]
13: 19058 Munich
    [pem_pos=0, final_scores=0.54141456,
final_local_scores=0.5355457, pem=-0.119245164,
similarity=20.719938]
```

Sample: b'692_Arch'

Arch [1 1 Alberta 1] well tests 1,100 bbl / day . [2 2 FORT W
ORTH 2] , [3 3 Texas 3] 1996-08-28 Arch Petroleum Inc said
Wednesday an exploratory well in [4 4 Alberta 4] 's [5 5 Mori
nville 5] area tested in excess of 1,100 barrels daily and
will begin production immediately . The company said the 90
percent owned Trax et al Morinville 10-23 logged 28 feet of
productive Leduc Reef at 5,350 feet . Reserve estimates from
this well are at 250,000 gross barrels of oil . The Trax well
is one of the prospects developed by Arch through its early
1996 purchase of Trax Petroleums Ltd , the company said . It
said another , the Cometra et al Morinville 11-13 , has
logged 128 feet of productive Leduc Reef at 5,400 feet and is
flowing water free at the rate of 590 barrels of oil per day
on a 15 / 64ths-inch choke . Arch owns a 16 percent working
interest in this well with most of the rest held by the
privately owned operator . Arch said full production has
begun and initial estimates of gross reserves attributable to
this well range up to one million barrels of oil . A third
well drilled in this area , the Trax et al Morinville 2-25 ,
encountered the Leduc Reef but tested wet . At the company 's
[6 6 Nordegg 6] prospect , the Apache et al Saunders 14-28 ,
has reached total depth in the Leduc Reef at approximately
3,800 feet and has been abandoned , Arch said . It said this
acreage earning well brought an interest in an additional
5,120 acres , building the company 's gross land position in
this area to 8,320 acres . At the Butte prospect , the
Garrington 4-8 has reached total depth of 11,500 and has
logged 278 feet of Leduc Reef . Testing has begun and results
will be announced within the next several days , Arch said .
Including the costs of the two abandoned wells , the company
said , these first prospects have added to reserves at a
finding cost of about [7 7 U.S. 7] \$ 2 per barrel of oil
equivalent . -- Jim Brumm 212-859-1710 .

1: 717 Alberta

2: 53839 Fort_Worth,_Texas

3: 29810 Texas

4: 717 Alberta

5: 2302768 Morinville

6: 5103900 Nordegg,_Alberta

7: 3434750 United_States

1: 717 Alberta

[pem_pos=0, final_scores=0.64027494,
final_local_scores=0.61366004, pem=-0.06637208,
similarity=22.763494]

2: 53839 Fort_Worth,_Texas

[pem_pos=0, final_scores=0.58403593,
final_local_scores=0.60446805, pem=0.0, similarity=22.128244]

3: 29810 Texas

[pem_pos=0, final_scores=0.79366666,
final_local_scores=0.81242466, pem=-0.16191302,
similarity=29.214594]

```
4: 717 Alberta
    [pem_pos=0, final_scores=0.64590204,
final_local_scores=0.6198068, pem=-0.06637208,
similarity=22.946926]
5: 2302768 Morinville
    [pem_pos=0, final_scores=0.40604812,
final_local_scores=0.4862315, pem=0.0, similarity=18.59983]
6: 5103900 Nordegg,_Alberta
    [pem_pos=0, final_scores=0.3920701,
final_local_scores=0.5093092, pem=0.0, similarity=19.288515]
7: 3434750 United_States
    [pem_pos=0, final_scores=0.3702938,
final_local_scores=0.36947268, pem=-0.14075686,
similarity=15.880971]
```

Sample: b'693_U.S.'

[1 1 U.S. 1] judge orders [2 2 Biogen 2] , [3 3 Berlex 3] officials deposed . Leslie Gevirtz [4 4 BOSTON 4] 1996-08-28 In order to help him decide whether he should hear the case , a [5 5 U.S. District Court 5] judge Wednesday ordered the legal counsels of Biogen Inc and [6 6 Berlex Laboratories 6] , a subsidiary of [7 7 Schering AG 7] , deposed . The tempest beyond the test tube involves allegations that the U.S. Food and Drug Administration violated the Orphan Drug law by allowing [8 8 Biogen 8] the right to sell its multiple sclerosis drug Avonex . [9 9 Berlex 9] also charges that Avonex is so similar to its MS drug , [10 10 Betaseron 10] , that it is a patent infringement . Both drugs are types of interferon . One analyst said sales of Avonex had already cut into [11 11 Betaseron 11] market share . BioVest Research , Inc 's analyst Eddie Hedaya said , " [12 12 Berlex 12] sales are losing share like mad ... my understanding of the marketplace is that they 're below expectations . " He added Chiron Corp reported its sales of inventory to [13 13 Berlex 13] was down . [14 14 Chiron 14] makes [15 15 Betaseron 15] ; [16 16 Berlex 16] markets it , he said . [17 17 Biogen 17] , in its Securities and Exchange Commission quarterly report for the period ending June 30 , said it had earned \$ 6.1 million from Avonex sales during the drugs first six weeks on the market . When it approved Avonex in May , the [18 18 FDA 18] said both [19 19 Biogen 19] 's product and [20 20 Betaseron 20] were developed under the incentives of the Ophran Drug Act which provides seven years of marketing exclusivity for products that treat rare diseases . Avonex " has been allowed to enter the market because it differs from interferon beta-1b ([21 21 Betaseron 21]) .. . " the [22 22 FDA 22] said . Now , [23 23 U.S. 23] District Judge [24 24 Mark Wolf 24] has ordered the chief counsel for [25 25 Biogen 25] , [26 26 Michael Astrue 26] , and Robert Chabora , his counterpart at [27 27 Berlex 27] be deposed about a May 21 meeting the two men attended to help him determine whether the lawsuit filed by [28 28 Biogen 28] against [29 29 Berlex 29] should

be heard in [30 30 Massachusetts 30] . [31 31 Berlex 31] filed a lawsuit against [32 32 Biogen 32] in [33 33 U.S. District Court 33] in [34 34 Newark 34] , [35 35 N.J. 35] in July , but [36 36 Biogen 36] had already filed a suit against [37 37 Berlex 37] in [38 38 Massachusetts 38] in May . [39 39 Wolf 39] ordered the depositions to determine if he or [40 40 U.S. 40] District Judge [41 41 John Bissell 41] of [42 42 Newark 42] should preside over the case .

1: 3434750 United_States
2: 380532 Biogen_Idec
3: 4383256 Bayer_HealthCare_Pharmaceuticals
4: 24437894 Boston
5: 70235 United_States_district_court
6: 4383256 Bayer_HealthCare_Pharmaceuticals
7: 2091870 Schering_AG
8: 380532 Biogen_Idec
9: 4383256 Bayer_HealthCare_Pharmaceuticals
10: 2858017 Interferon_beta_1b
11: 2858017 Interferon_beta_1b
12: 4383256 Bayer_HealthCare_Pharmaceuticals
13: 4383256 Bayer_HealthCare_Pharmaceuticals
14: 1184467 Chiron_Corporation
15: 2858017 Interferon_beta_1b
16: 4383256 Bayer_HealthCare_Pharmaceuticals
17: 380532 Biogen_Idec
18: 3709121 FDA_(trade_union)
19: 380532 Biogen_Idec
20: 2858017 Interferon_beta_1b
21: 2858017 Interferon_beta_1b
22: 3709121 FDA_(trade_union)
23: 3434750 United_States
24: 21835440 Mark_L._Wolf
25: 380532 Biogen_Idec
26: 11076522 Michael_J._Astrue
27: 4383256 Bayer_HealthCare_Pharmaceuticals
28: 380532 Biogen_Idec
29: 4383256 Bayer_HealthCare_Pharmaceuticals
30: 1645518 Massachusetts
31: 4383256 Bayer_HealthCare_Pharmaceuticals
32: 380532 Biogen_Idec
33: 70235 United_States_district_court
34: 67234 Newark,_New_Jersey
35: 21648 New_Jersey
36: 380532 Biogen_Idec
37: 4383256 Bayer_HealthCare_Pharmaceuticals
38: 1645518 Massachusetts
39: 21835440 Mark_L._Wolf
40: 3434750 United_States
41: 21804225 John_Winslow_Bissell
42: 67234 Newark,_New_Jersey
1: 3434750 United_States
[pem_pos=0, final_scores=0.55553335,

```
final_local_scores=0.611201, pem=-0.14075686,
similarity=23.094625]
2: 380532 Biogen_Idec
    [pem_pos=0, final_scores=0.7092404,
final_local_scores=0.7414939, pem=0.0, similarity=26.21737]
3: 4383256 Bayer_HealthCare_Pharmaceuticals
    [pem_pos=0, final_scores=0.6982053,
final_local_scores=0.73028207, pem=0.0, similarity=25.882786]
4: 24437894 Boston
    [pem_pos=0, final_scores=0.56970567,
final_local_scores=0.63469386, pem=-0.23164093,
similarity=24.289944]
5: 70235 United_States_district_court
    [pem_pos=0, final_scores=0.28400618,
final_local_scores=0.3466847, pem=-0.01716152,
similarity=14.528803]
6: 4383256 Bayer_HealthCare_Pharmaceuticals
    [pem_pos=0, final_scores=0.49952894,
final_local_scores=0.51325834, pem=0.0, similarity=19.406366]
7: 2091870 Schering_AG
    [pem_pos=0, final_scores=0.33669806,
final_local_scores=0.3876062, pem=0.0, similarity=15.656655]
8: 380532 Biogen_Idec
    [pem_pos=0, final_scores=0.6298507,
final_local_scores=0.6547727, pem=0.0, similarity=23.629436]
9: 4383256 Bayer_HealthCare_Pharmaceuticals
    [pem_pos=0, final_scores=0.64498717,
final_local_scores=0.6721493, pem=0.0, similarity=24.14799]
10: 2858017 Interferon_beta_1b
    [pem_pos=0, final_scores=0.6173436,
final_local_scores=0.7069604, pem=0.0, similarity=25.186821]
11: 2858017 Interferon_beta_1b
    [pem_pos=0, final_scores=0.53786695,
final_local_scores=0.6201442, pem=0.0, similarity=22.596052]
12: 4383256 Bayer_HealthCare_Pharmaceuticals
    [pem_pos=0, final_scores=0.767374,
final_local_scores=0.8058384, pem=0.0, similarity=28.137539]
13: 4383256 Bayer_HealthCare_Pharmaceuticals
    [pem_pos=0, final_scores=0.777361,
final_local_scores=0.8167476, pem=0.0, similarity=28.46309]
14: 1184467 Chiron_Corporation
    [pem_pos=3, final_scores=0.38510782,
final_local_scores=0.42554843, pem=-2.0596766,
similarity=27.989769]
15: 2858017 Interferon_beta_1b
    [pem_pos=0, final_scores=0.6412412,
final_local_scores=0.7330649, pem=0.0, similarity=25.965832]
16: 4383256 Bayer_HealthCare_Pharmaceuticals
    [pem_pos=0, final_scores=0.6909477,
final_local_scores=0.72235423, pem=0.0, similarity=25.646204]
17: 380532 Biogen_Idec
    [pem_pos=0, final_scores=0.6245384,
```

```
final_local_scores=0.6489698, pem=0.0, similarity=23.456266]
18: 3709121 FDA_(trade_union)
    [pem_pos=0, final_scores=0.3639726,
final_local_scores=0.50163114, pem=0.0, similarity=19.059387]
19: 380532 Biogen_Idec
    [pem_pos=0, final_scores=0.65044075,
final_local_scores=0.6772642, pem=0.0, similarity=24.300629]
20: 2858017 Interferon_beta_1b
    [pem_pos=0, final_scores=0.62877196,
final_local_scores=0.7194441, pem=0.0, similarity=25.55936]
21: 2858017 Interferon_beta_1b
    [pem_pos=0, final_scores=0.5557544,
final_local_scores=0.63968354, pem=0.0, similarity=23.179146]
22: 3709121 FDA_(trade_union)
    [pem_pos=0, final_scores=0.45625883,
final_local_scores=0.6024398, pem=0.0, similarity=22.067719]
23: 3434750 United_States
    [pem_pos=0, final_scores=0.5309289,
final_local_scores=0.5843243, pem=-0.14075686,
similarity=22.292574]
24: 21835440 Mark_L._Wolf
    [pem_pos=0, final_scores=0.5618699,
final_local_scores=0.7105099, pem=0.0, similarity=25.292746]
25: 380532 Biogen_Idec
    [pem_pos=0, final_scores=0.6327052,
final_local_scores=0.65789086, pem=0.0, similarity=23.722488]
26: 11076522 Michael_J._Astrue
    [pem_pos=0, final_scores=0.40625232,
final_local_scores=0.5522325, pem=0.0, similarity=20.569433]
27: 4383256 Bayer_HealthCare_Pharmaceuticals
    [pem_pos=0, final_scores=0.6320556,
final_local_scores=0.65802354, pem=0.0, similarity=23.726448]
28: 380532 Biogen_Idec
    [pem_pos=0, final_scores=0.76395124,
final_local_scores=0.8012572, pem=0.0, similarity=28.000826]
29: 4383256 Bayer_HealthCare_Pharmaceuticals
    [pem_pos=0, final_scores=0.7433011,
final_local_scores=0.7795424, pem=0.0, similarity=27.352812]
30: 1645518 Massachusetts
    [pem_pos=0, final_scores=0.48459488,
final_local_scores=0.5383997, pem=-0.1423749,
similarity=20.93089]
31: 4383256 Bayer_HealthCare_Pharmaceuticals
    [pem_pos=0, final_scores=0.6955989,
final_local_scores=0.727435, pem=0.0, similarity=25.797825]
32: 380532 Biogen_Idec
    [pem_pos=0, final_scores=0.68931586,
final_local_scores=0.71972936, pem=0.0, similarity=25.567873]
33: 70235 United_States_district_court
    [pem_pos=0, final_scores=0.43124348,
final_local_scores=0.50751907, pem=-0.01716152,
similarity=19.328423]
```

```
34: 67234 Newark,_New_Jersey
    [pem_pos=0, final_scores=0.37376702,
final_local_scores=0.42185682, pem=-0.6286074,
similarity=20.097227]
35: 21648 New_Jersey
    [pem_pos=0, final_scores=0.46823418,
final_local_scores=0.5188206, pem=0.0, similarity=19.572353]
36: 380532 Biogen_Idec
    [pem_pos=0, final_scores=0.65296024,
final_local_scores=0.6800164, pem=0.0, similarity=24.38276]
37: 4383256 Bayer_HealthCare_Pharmaceuticals
    [pem_pos=0, final_scores=0.67792827,
final_local_scores=0.7081325, pem=0.0, similarity=25.2218]
38: 1645518 Massachusetts
    [pem_pos=0, final_scores=0.5838371,
final_local_scores=0.6468067, pem=-0.1423749,
similarity=24.165972]
39: 21835440 Mark_L._Wolf
    [pem_pos=0, final_scores=0.6366392,
final_local_scores=0.79218394, pem=0.0, similarity=27.730062]
40: 3434750 United_States
    [pem_pos=0, final_scores=0.5589178,
final_local_scores=0.6148979, pem=-0.14075686,
similarity=23.20495]
41: 21804225 John_Winslow_Bissell
    [pem_pos=0, final_scores=0.5437438,
final_local_scores=0.70582753, pem=0.0, similarity=25.153015]
42: 67234 Newark,_New_Jersey
    [pem_pos=0, final_scores=0.36325562,
final_local_scores=0.4103747, pem=-0.6286074,
similarity=19.754578]
```

Sample: b'694_U.S.'

```
[1 1 U.S. Treasury 1] balances at [2 2 Fed 2] rose on Aug
27 . [3 3 WASHINGTON 3] 1996-08-28 [4 4 U.S. Treasury 4]
balances at [5 5 Federal Reserve 5] based on [6 6 Treasury De
partment 6] 's latest budget statement . ( BILLIONS OF DLRS )
Aug 27 Aug 26 [7 7 Fed 7] acct 5.208 4.425 Tax / loan note
acct 14.828 15.687 Cash balance 20.036 20.112 --- Total
public debt , subject to limit 5,124.053 5,122.084
1: 53667 United_States_Department_of_the_Treasury
2: 10819 Federal_Reserve_System
3: 108956 Washington,_D.C.
4: 53667 United_States_Department_of_the_Treasury
5: 10819 Federal_Reserve_System
6: 53667 United_States_Department_of_the_Treasury
7: 10819 Federal_Reserve_System
1: 53667 United_States_Department_of_the_Treasury
    [pem_pos=0, final_scores=0.47065824,
final_local_scores=0.41682434, pem=0.0, similarity=16.528584]
2: 10819 Federal_Reserve_System
    [pem_pos=0, final_scores=0.44100082,
```

```
final_local_scores=0.39453256, pem=0.0, similarity=15.863352]
3: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.5165805,
final_local_scores=0.5699608, pem=-1.7496481,
similarity=30.613338]
4: 53667 United_States_Department_of_the_Treasury
    [pem_pos=0, final_scores=0.43550545,
final_local_scores=0.37842524, pem=0.0, similarity=15.382677]
5: 10819 Federal_Reserve_System
    [pem_pos=0, final_scores=0.41192204,
final_local_scores=0.36276838, pem=0.0, similarity=14.915445]
6: 53667 United_States_Department_of_the_Treasury
    [pem_pos=0, final_scores=0.37917358,
final_local_scores=0.3168912, pem=0.0, similarity=13.546378]
7: 10819 Federal_Reserve_System
    [pem_pos=0, final_scores=0.51006454,
final_local_scores=0.46997422, pem=0.0, similarity=18.114681]
```

Sample: b'695_COMEX'

[1 1 COMEX 1] copper ends higher after late recovery . [2 2 N
EW YORK 2] 1996-08-28 [3 3 COMEX 3] copper ended higher after
a late , modest recovery dragged the market from the lows ,
but traders shrugged-off an imminent strike at [4 4 Codelco 4]
's Salvador mine in [5 5 Chile 5] . The market was also
waiting for Friday 's LME stock report which will include
figures delayed from Tuesday because of the [6 6 U.K. 6]
public holiday on Monday when the LME was closed . " We are
in the late stages of the weaker period of the market , and
as we get to the post [7 7 Labor Day 7] market we will start
to see more consumer interest and the demand side of the
market will start to firm prices up , " said William O'Neill
of [8 8 Merrill Lynch 8] . December [9 9 COMEX 9] settled
0.35 cent higher at 90.20 cents , traded 90.50 to 89.40
cents . September went out 0.05 cent lower at 91.05 . The
August contract expired at 0.85 cent down at 90.85 cents .
Volume was estimated at 8,000 lots . Workers at Salvador
voted to strike from Saturday , and it was not clear when
further talks between the unions and management would take
place . " Salvador is a small facility , and the prospects
are that if there will be a strike , it will not be a long
strike , " O'Neill said . -- Huw Jones , New York Commodities
212-859-1646

```
1: 469729 New_York_Mercantile_Exchange
2: 8210131 New_York
3: 469729 New_York_Mercantile_Exchange
4: 985804 Codelco
5: 5489 Chile
6: 31717 United_Kingdom
7: 19156272 Labor_Day
8: 33731299 Merrill_Lynch
9: 469729 New_York_Mercantile_Exchange
1: 469729 New_York_Mercantile_Exchange
```

```
[pem_pos=0, final_scores=0.64939934,
final_local_scores=0.6802763, pem=0.0, similarity=24.390514]
2: 8210131 New_York
[pem_pos=0, final_scores=0.20763232,
final_local_scores=0.2017254, pem=-0.2889045,
similarity=11.680708]
3: 469729 New_York_Mercantile_Exchange
[pem_pos=0, final_scores=0.5848964,
final_local_scores=0.6098166, pem=0.0, similarity=22.287857]
4: 985804 Codelco
[pem_pos=0, final_scores=0.411138,
final_local_scores=0.5177355, pem=0.0, similarity=19.539972]
5: 5489 Chile
[pem_pos=0, final_scores=0.63862425,
final_local_scores=0.72183573, pem=-0.27404457,
similarity=27.121027]
6: 31717 United_Kingdom
[pem_pos=0, final_scores=0.3043213,
final_local_scores=0.3197427, pem=-0.13569236,
similarity=14.36939]
7: 19156272 Labor_Day
[pem_pos=0, final_scores=0.3551038,
final_local_scores=0.3984143, pem=0.0, similarity=15.979191]
8: 33731299 Merrill_Lynch
[pem_pos=0, final_scores=0.45630842,
final_local_scores=0.5076331, pem=0.0, similarity=19.238497]
9: 469729 New_York_Mercantile_Exchange
[pem_pos=0, final_scores=0.63504773,
final_local_scores=0.66459936, pem=0.0, similarity=23.922684]
```

Sample: b'696_U.S.'

```
[1 1 U.S. 1] copper service center shipments stable - CBSA .
[2 2 NEW YORK 2] 1996-08-28 Average daily shipments from [3 3
U.S. 3] copper service centers in July fell three percent
from the previous month , but were higher than in July 1995 ,
the Copper and Brass Servicenter Association reported . "
July was still above the historic average for that month , "
the CBSA said . In the first seven months of 1996 , shipments
of copper sheet , coil and strip were 2.2 percent ahead of
the same period last year . Alloy shipments , however , were
7.5 percent down . " Several service centers indicated that
while their volume of orders remains constant , the size-per-
order continues to be smaller than what was realised during
the first five months of the year , " the CBSA said . Service
centers continued to lower their inventories in July when
total copper stocks were off two percent and alloy products
down 1.9 percent . -- New York Commodities 212-859-1646
1: 3434750 United_States
2: 8210131 New_York
3: 3434750 United_States
1: 3434750 United_States
[pem_pos=0, final_scores=0.5142798,
```

```
final_local_scores=0.46327525, pem=-0.14075686,
similarity=18.680227]
2: 8210131 New_York
    [pem_pos=0, final_scores=0.2804854,
final_local_scores=0.24334133, pem=-0.2889045,
similarity=12.922611]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.5637814,
final_local_scores=0.51734823, pem=-0.14075686,
similarity=20.293875]

Sample: b'697_Harleysville'
    Harleysville Group ups qrtly dividend . [1 1 HARLEYSVILLE 1]
    , Pa . 1996-08-28 Quarterly Latest Prior Amount $ 0.21 $ 0.19
    Pay Sept 30 Record Sept 16
1: 133365 Harleysville,_Pennsylvania
1: 133365 Harleysville,_Pennsylvania
    [pem_pos=0, final_scores=0.25749344,
final_local_scores=0.55385965, pem=0.0, similarity=20.61799]

Sample: b'698_Chile'
    [1 1 Chile 1] 's ENAP buys Oriente , Escravos crude for
Oct . [2 2 NEW YORK 2] 1996-08-28 [3 3 Chile 3] 's state oil
company Empresa Nacional del Petroleo ( ENAP ) bought a
second spot cargo of Oriente and nearly one million barrels
of Escravos in a recent tender , traders said Wednesday . A
400,000 barrel cargo of Ecuadorian Oriente and 960,000
barrels of [4 4 Nigerian 4] Escravos was awarded in a tender
for Oct 15-18 late last week , but price information remains
vague . " The Oriente will be supplied by the same seller at
a small premium to formula , " a trade source said ,
referring to the first October cargo sold two weeks ago at
Petroecuador 's sale formula plus five cents fob . Escravos
was sold on a Dated Brent related basis , with premiums for
the light grade seen in the low 50-cent range . The next
purchase tender from ENAP is expected for late October or
early November crude , traders said . -- Jacqueline Wong ,
New York Energy Desk +1 212 859 1620
1: 5489 Chile
2: 8210131 New_York
3: 5489 Chile
4: 21383 Nigeria
1: 5489 Chile
    [pem_pos=0, final_scores=0.61341435,
final_local_scores=0.60370964, pem=-0.27404457,
similarity=23.59591]
2: 8210131 New_York
    [pem_pos=0, final_scores=0.3125404,
final_local_scores=0.33543223, pem=-0.2889045,
similarity=15.670788]
3: 5489 Chile
    [pem_pos=0, final_scores=0.60949,
```

```
final_local_scores=0.59942293, pem=-0.27404457,  
similarity=23.467985]  
4: 21383 Nigeria  
    [pem_pos=0, final_scores=0.5031977,  
final_local_scores=0.53045493, pem=-0.075027294,  
similarity=20.327557]
```

Sample: b'699_Reuters'

Reuters historical calendar - September 4 . [1 1 LONDON 1] 1996-08-28 Following are some of the major events to have occurred on September 4 in history . 1241 - [2 2 Alexander III 2] , King of [3 3 Scotland 3] , born . King from 1249-1286 , he consolidated royal power , leaving [4 4 Scotland 4] united and independent . 1260 - The Ghibellines retook the city of [5 5 Florence 5] from the Florentine Guelfs at the battle of Monte Aperto . 1768 - Francois-Rene (Vicomte de) Chateaubriand born . He was a politician , one of the first [6 6 French 6] romantic writers and ambassador to the [7 7 British 7] court . He wrote " [8 8 Rene 8] " , a seminal work in the [9 9 French 9] romantic movement and a famous autobiography " Memoires d'Outre Tombe " . 1781 - [10 10 Los Angeles 10] was founded by [11 11 Spanish 11] settlers and named " El Pueblo de Nuestra Senora La Reina de Los Angeles " (The Town of Our Lady the Queen of the Angels) . 1824 - [1 2 12 Anton Bruckner 12] born . [13 13 Austrian 13] composer and organist , he wrote nine symphonies on a huge scale and three grand masses in the romantic tradition . 1870 - In [14 1 4 France 14] , the [15 15 Second Empire 15] was ended and [16 16 Napoleon III 16] was deposed after his surrender two days earlier in the Franco-Prussian war . 1886 - At [17 17 Skeleton Canyon 17] in [18 18 Arizona 18] , [19 19 Geronimo 19] , Apache chief and leader of the last great [20 20 Red Indian 20] rebellion finally surrendered to General [21 21 Nelson Miles 21] . 1892 - Prolific [22 22 French 22] modernist composer [23 23 Darius Milhaud 23] born . He wrote a jazz ballet " [24 24 La Creation du Monde 24] " and scores for many films including an early version of " [25 25 Madame Bovary 25] " . 1906 - German-born [26 26 U.S. 26] biologist [27 27 Max Delbrück 27] born . Winner of the 1969 [28 28 Nobel Prize 28] for physiology or medicine for work on the genetic structure of viruses that infect bacteria . 1907 - [29 29 Edvard Grieg 29] , [30 30 Norwegian 30] composer best known for his " [31 31 Peer Gynt Suite 31] " and his Piano Concerto , died in [32 32 Bergen 32] . 1908 - [33 33 U.S. 33] film director [34 34 Edward Dmytryk 34] born . Best known for his films " [35 35 Crossfire 35] " - one of [36 36 Hollywood 36] 's first attempts to deal with racial discrimination and " Farewell My lovely " . 1909 - The world 's first Boy Scout Rally was held at [37 37 Crystal Palace 37] near [38 38 London 38] . 1944 - [39 39 Brussels 39] and [40 40 Antwerp 40] in [41 41 Belgium 41] were liberated by [42 42 British 42] and [43 43 Canadian 43] troops in [44 44 World War Two 44] . 1948 - [45 45 Wilhelmina

45] , Queen of the [46 46 Netherlands 46] from 1890 and throughout World Wars One and [47 47 Two 47] abdicated in favour of her daughter [48 48 Juliana 48] . 1963 - [49 49 Robert Schuman 49] , [50 50 French 50] statesman , Prime Minister 1947-48 and Foreign Minister 1948-52 , died . He was responsible for the establishment of the European Coal and Steel Community . 1964 - The [51 51 Forth Road Bridge 51] in [52 52 Scotland 52] , measuring 6156 ft , and with a centre span of 3300 ft , was opened by Her Majesty the Queen . 1965 - [53 53 Albert Schweitzer 53] , theologian , philosopher and organist died in [54 54 Gabon 54] where he had set up a hospital in 1913 . Acclaimed for his interpretations of [55 55 J.S. Bach 55] 's works , he also won the [56 56 Nobel Peace Prize 56] for his efforts on behalf of the " Brotherhood of Nations " in 1952 . 1972 - At the 1 Olympic Games , [57 57 U.S. 57] swimmer [58 58 Mark Spitz 58] won his seventh gold medal , a record for a single [59 59 Olympiad 59] . 1974 - [60 60 East Germany 60] and the [61 61 United States 61] established formal diplomatic relations for the first time . 1977 - E.F. (Fritz) Schumacher , economic guru and author of the best seller " [62 62 Small is Beautiful 62] " , died on his way to a conference in [63 63 Switzerland 63] . 1989 - [64 64 Georges Simenon 64] , writer of 84 books based on the detective character [65 65 Inspector Maigret 65] , died . 1992 - [66 66 Bulgaria 66] 's former [1 2 Communist 1] leader [67 67 Todor Zhivkov 67] , deposed in 1989 , was sentenced to seven years in prison after being found guilty of embezzling state funds . 1995 - Declaring " united [68 68 Jerusalem 68] is ours " , [69 69 Israel 69] launched a 15-month celebration of the 3,000th anniversary of King [70 70 David 70] 's proclamation of the city as the capital of the [71 71 Jewish 71] people . 1995 - The [72 72 Fourth World Conference on Women 72] , the biggest [73 73 U.N. 73] gathering in history , began in [74 74 China 74] 's [75 75 Great Hall of the People 75] with a [76 76 U.N. 76] declaration that sexual equality was the last great project of the 20th century .

- 1: 17867 London
- 2: 1596 Alexander_III_of_Scotland
- 3: 26994 Scotland
- 4: 26994 Scotland
- 5: 11525 Florence
- 6: 5843419 France
- 7: 31717 United_Kingdom
- 8: 5963679 René_(novella)
- 9: 5843419 France
- 10: 18110 Los_Angeles
- 11: 26667 Spain
- 12: 44012 Anton_Bruckner
- 13: 26964606 Austria
- 14: 5843419 France
- 15: 216066 Second_French_Empire
- 16: 62581 Napoleon_III

17: 2702049 Skeleton_Canyon
18: 21883824 Arizona
19: 221765 Geronimo
20: 21217 Native_Americans_in_the_United_States
21: 1291705 Nelson_A._Miles
22: 5843419 France
23: 85267 Darius_Milhaud
24: 1968469 La_création_du_monde
25: 161242 Madame_Bovary
26: 3434750 United_States
27: 169382 Max_Delbrück
28: 21201 Nobel_Prize
29: 9514 Edvard_Grieg
30: 21241 Norway
31: 555554 Peer_Gynt
32: 56494 Bergen
33: 3434750 United_States
34: 382196 Edward_Dmytryk
35: 42106 Crossfire_(film)
36: 53849 Hollywood
37: 93961 Crystal_Palace,_London
38: 17867 London
39: 3708 Brussels
40: 32149462 Antwerp
41: 3343 Belgium
42: 31717 United_Kingdom
43: 5042916 Canada
44: 32927 World_War_II
45: 102298 Wilhelmina_of_the_Netherlands
46: 21148 Netherlands
47: 32927 World_War_II
48: 80481 Juliana_of_the_Netherlands
49: 374183 Robert_Schuman
50: 5843419 France
51: 144607 Forth_Road_Bridge
52: 26994 Scotland
53: 1029 Albert_Schweitzer
54: 12027 Gabon
55: 9906294 Johann_Sebastian_Bach
56: 26230922 Nobel_Peace_Prize
57: 3434750 United_States
58: 66265 Mark_Spitz
59: 332902 Olympiad
60: 13058 East_Germany
61: 3434750 United_States
62: 2514850 Small_Is_Beautiful
63: 26748 Switzerland
64: 182032 Georges_Simenon
65: 69445 Jules_Maigret
66: 3415 Bulgaria
67: 276706 Todor_Zhivkov
68: 16043 Jerusalem

69: 9282173 Israel
70: 8551 David
71: 25955086 Jews
72: 3523938 Fourth_World_Conference_on_Women
73: 31769 United_Nations
74: 5405 China
75: 553623 Great_Hall_of_the_People
76: 31769 United_Nations
1: 83112 1972_Summer_Olympics
 [winner: 22576 Olympic_Games, pem_pos=0,
final_scores=-0.11118546, final_local_scores=-0.14950904,
pem=-0.5587218, similarity=2.666481]
 RECALL MISS: 83112 1972_Summer_Olympics]
2: 37008 Communist_party
 [winner: 9209651 Communism, pem_pos=0,
final_scores=0.049156323, final_local_scores=0.02589716,
pem=-0.39361537, similarity=7.0030775]
 gt: 37008 Communist_party, pem_pos=9,
final_scores=-0.08673607, final_local_scores=-0.10503733,
pem=-3.6379085, similarity=20.7387]
1: 17867 London
 [pem_pos=0, final_scores=0.6874643,
final_local_scores=0.6807073, pem=-0.09086737,
similarity=24.897526]
2: 1596 Alexander_III_of_Scotland
 [pem_pos=1, final_scores=0.36376053,
final_local_scores=0.43614948, pem=-1.2571235,
similarity=23.941715]
3: 26994 Scotland
 [pem_pos=0, final_scores=0.42774636,
final_local_scores=0.42969728, pem=-0.22905664,
similarity=18.158382]
4: 26994 Scotland
 [pem_pos=3, final_scores=0.42433125,
final_local_scores=0.4259668, pem=-0.22905664,
similarity=18.047058]
5: 11525 Florence
 [pem_pos=0, final_scores=0.5202921,
final_local_scores=0.53604084, pem=-0.32229787,
similarity=21.838947]
6: 5843419 France
 [pem_pos=0, final_scores=0.73658913,
final_local_scores=0.7297004, pem=-0.80223274,
similarity=30.228094]
7: 31717 United_Kingdom
 [pem_pos=0, final_scores=0.5955637,
final_local_scores=0.5808806, pem=-0.48936236,
similarity=24.085577]
8: 5963679 René_(novella)
 [pem_pos=8, final_scores=0.5754794,
final_local_scores=0.74282306, pem=-3.6826568,
similarity=46.283894]

```
9: 5843419 France
    [pem_pos=0, final_scores=0.7654682,
final_local_scores=0.7612464, pem=-0.80223274,
similarity=31.169489]
10: 18110 Los_Angeles
    [pem_pos=0, final_scores=0.5469493,
final_local_scores=0.55927086, pem=-0.100652285,
similarity=21.326834]
11: 26667 Spain
    [pem_pos=1, final_scores=0.5152429,
final_local_scores=0.5217976, pem=-1.2425895,
similarity=26.418589]
12: 44012 Anton_Bruckner
    [pem_pos=0, final_scores=0.38400155,
final_local_scores=0.462659, pem=0.0, similarity=17.89638]
13: 26964606 Austria
    [pem_pos=0, final_scores=0.6788762,
final_local_scores=0.67509055, pem=-0.40674555,
similarity=26.447706]
14: 5843419 France
    [pem_pos=0, final_scores=0.70844656,
final_local_scores=0.6989589, pem=-0.16240834,
similarity=25.831242]
15: 216066 Second_French_Empire
    [pem_pos=0, final_scores=0.30580133,
final_local_scores=0.31449163, pem=-0.011029043,
similarity=13.534748]
16: 62581 Napoleon_III
    [pem_pos=0, final_scores=0.4457435,
final_local_scores=0.46822137, pem=0.0, similarity=18.062372]
17: 2702049 Skeleton_Canyon
    [pem_pos=0, final_scores=0.4305166,
final_local_scores=0.5722107, pem=0.0, similarity=21.165623]
18: 21883824 Arizona
    [pem_pos=0, final_scores=0.6118656,
final_local_scores=0.6778077, pem=-0.22599088,
similarity=25.54582]
19: 221765 Geronimo
    [pem_pos=0, final_scores=0.38604754,
final_local_scores=0.4607544, pem=-0.10300582,
similarity=18.399706]
20: 21217 Native_Americans_in_the_United_States
    [pem_pos=1, final_scores=0.16149323,
final_local_scores=0.15950158, pem=-0.8484681,
similarity=13.463659]
21: 1291705 Nelson_A._Miles
    [pem_pos=0, final_scores=0.3924576,
final_local_scores=0.47584718, pem=0.0, similarity=18.289942]
22: 5843419 France
    [pem_pos=0, final_scores=0.81635803,
final_local_scores=0.8168358, pem=-0.80223274,
similarity=32.82839]
```

```
23: 85267 Darius_Milhaud
    [pem_pos=0, final_scores=0.45532733,
final_local_scores=0.52714074, pem=0.0, similarity=19.820644]
24: 1968469 La_création_du_monde
    [pem_pos=0, final_scores=0.35730654,
final_local_scores=0.5053423, pem=0.0, similarity=19.170135]
25: 161242 Madame_Bovary
    [pem_pos=0, final_scores=0.41485828,
final_local_scores=0.5105877, pem=0.0, similarity=19.326668]
26: 3434750 United_States
    [pem_pos=0, final_scores=0.51432276,
final_local_scores=0.4920209, pem=-0.14075686,
similarity=19.538055]
27: 169382 Max_Delbrück
    [pem_pos=0, final_scores=0.5226298,
final_local_scores=0.6552346, pem=0.0, similarity=23.643219]
28: 21201 Nobel_Prize
    [pem_pos=0, final_scores=0.22260408,
final_local_scores=0.24801856, pem=-0.15178943,
similarity=12.316536]
29: 9514 Edvard_Grieg
    [pem_pos=0, final_scores=0.35300684,
final_local_scores=0.3983329, pem=0.0, similarity=15.976761]
30: 21241 Norway
    [pem_pos=0, final_scores=0.6130947,
final_local_scores=0.64588773, pem=-0.605967,
similarity=26.659634]
31: 555554 Peer_Gynt
    [pem_pos=1, final_scores=0.3808524,
final_local_scores=0.44166, pem=-0.4677589,
similarity=19.813475]
32: 56494 Bergen
    [pem_pos=0, final_scores=0.42562646,
final_local_scores=0.44778955, pem=-0.28360042,
similarity=18.99491]
33: 3434750 United_States
    [pem_pos=0, final_scores=0.5067796,
final_local_scores=0.48378116, pem=-0.14075686,
similarity=19.292166]
34: 382196 Edward_Dmytryk
    [pem_pos=0, final_scores=0.51570684,
final_local_scores=0.6196538, pem=0.0, similarity=22.581419]
35: 42106 Crossfire_(film)
    [pem_pos=1, final_scores=0.46690673,
final_local_scores=0.6014077, pem=-1.798317,
similarity=31.816448]
36: 53849 Hollywood
    [pem_pos=1, final_scores=0.44693977,
final_local_scores=0.45322603, pem=-2.0106025,
similarity=28.54885]
37: 93961 Crystal_Palace,_London
    [pem_pos=2, final_scores=0.25676662,
```

```
final_local_scores=0.30270317, pem=-2.1751254,  
similarity=24.951649]  
38: 17867 London  
    [pem_pos=0, final_scores=0.6152372,  
final_local_scores=0.60181004, pem=-0.09086737,  
similarity=22.543076]  
39: 3708 Brussels  
    [pem_pos=0, final_scores=0.6815339,  
final_local_scores=0.6733486, pem=-0.025440974,  
similarity=24.32213]  
40: 32149462 Antwerp  
    [pem_pos=0, final_scores=0.5872442,  
final_local_scores=0.59502906, pem=-0.11904648,  
similarity=22.49396]  
41: 3343 Belgium  
    [pem_pos=0, final_scores=0.62062436,  
final_local_scores=0.6052343, pem=-0.24180195,  
similarity=23.46607]  
42: 31717 United_Kingdom  
    [pem_pos=0, final_scores=0.59085006,  
final_local_scores=0.5757316, pem=-0.48936236,  
similarity=23.931921]  
43: 5042916 Canada  
    [pem_pos=0, final_scores=0.53814405,  
final_local_scores=0.54015344, pem=-0.41303828,  
similarity=22.455135]  
44: 32927 World_War_II  
    [pem_pos=0, final_scores=0.5300203,  
final_local_scores=0.5168522, pem=0.0, similarity=19.513613]  
45: 102298 Wilhelmina_of_the_Netherlands  
    [pem_pos=0, final_scores=0.4193321,  
final_local_scores=0.46831387, pem=-0.35392305,  
similarity=19.989822]  
46: 21148 Netherlands  
    [pem_pos=0, final_scores=0.43500155,  
final_local_scores=0.41536856, pem=-0.24940795,  
similarity=17.84146]  
47: 32927 World_War_II  
    [pem_pos=15, final_scores=0.71928424,  
final_local_scores=0.72359437, pem=-3.6182806,  
similarity=45.359985]  
48: 80481 Juliana_of_the_Netherlands  
    [pem_pos=0, final_scores=0.56747764,  
final_local_scores=0.6489691, pem=-0.5059484,  
similarity=26.207672]  
49: 374183 Robert_Schuman  
    [pem_pos=0, final_scores=0.50977963,  
final_local_scores=0.57547647, pem=0.0, similarity=21.263079]  
50: 5843419 France  
    [pem_pos=0, final_scores=0.7201244,  
final_local_scores=0.71171516, pem=-0.80223274,  
similarity=29.69138]
```

```
51: 144607 Forth_Road_Bridge
    [pem_pos=0, final_scores=0.5000101,
final_local_scores=0.63760453, pem=0.0, similarity=23.117104]
52: 26994 Scotland
    [pem_pos=3, final_scores=0.46766728,
final_local_scores=0.4733048, pem=-0.22905664,
similarity=19.459719]
53: 1029 Albert_Schweitzer
    [pem_pos=0, final_scores=0.45014435,
final_local_scores=0.4865172, pem=0.0, similarity=18.608356]
54: 12027 Gabon
    [pem_pos=0, final_scores=0.44950426,
final_local_scores=0.5191935, pem=-0.3021721,
similarity=21.226742]
55: 9906294 Johann_Sebastian_Bach
    [pem_pos=0, final_scores=0.4433505,
final_local_scores=0.49024886, pem=0.0, similarity=18.719717]
56: 26230922 Nobel_Peace_Prize
    [pem_pos=0, final_scores=0.35814214,
final_local_scores=0.38691872, pem=0.0, similarity=15.636139]
57: 3434750 United_States
    [pem_pos=0, final_scores=0.59282845,
final_local_scores=0.5777765, pem=-0.14075686,
similarity=22.097174]
58: 66265 Mark_Spitz
    [pem_pos=0, final_scores=0.5440691,
final_local_scores=0.6679841, pem=0.0, similarity=24.023691]
59: 332902 Olympiad
    [pem_pos=0, final_scores=0.3740666,
final_local_scores=0.42253518, pem=0.0, similarity=16.699005]
60: 13058 East_Germany
    [pem_pos=0, final_scores=0.3699959,
final_local_scores=0.36798912, pem=-0.26717973,
similarity=16.524206]
61: 3434750 United_States
    [pem_pos=0, final_scores=0.4351645,
final_local_scores=0.40555257, pem=-0.073457584,
similarity=16.591684]
62: 2514850 Small_Is_Beautiful
    [pem_pos=0, final_scores=0.6346412,
final_local_scores=0.78590435, pem=0.0, similarity=27.542665]
63: 26748 Switzerland
    [pem_pos=0, final_scores=0.65533906,
final_local_scores=0.6546431, pem=-0.18598008,
similarity=24.636957]
64: 182032 Georges_Simenon
    [pem_pos=0, final_scores=0.44007128,
final_local_scores=0.5094568, pem=0.0, similarity=19.292921]
65: 69445 Jules_Maigret
    [pem_pos=0, final_scores=0.50487345,
final_local_scores=0.61687964, pem=0.0, similarity=22.498632]
66: 3415 Bulgaria
```

```
[pem_pos=0, final_scores=0.69485176,
final_local_scores=0.73296946, pem=-0.31727514,
similarity=27.688375]
67: 276706 Todor_Zhivkov
    [pem_pos=0, final_scores=0.6034274,
final_local_scores=0.7129824, pem=0.0, similarity=25.36653]
68: 16043 Jerusalem
    [pem_pos=0, final_scores=0.5496833,
final_local_scores=0.56883264, pem=-0.13239029,
similarity=21.784773]
69: 9282173 Israel
    [pem_pos=0, final_scores=0.53165096,
final_local_scores=0.5403419, pem=-0.21279564,
similarity=21.371809]
70: 8551 David
    [pem_pos=0, final_scores=0.4396026,
final_local_scores=0.5034445, pem=-0.8369082,
similarity=23.664738]
71: 25955086 Jews
    [pem_pos=1, final_scores=0.456114,
final_local_scores=0.45637417, pem=-0.80476934,
similarity=22.085289]
72: 3523938 Fourth_World_Conference_on_Women
    [pem_pos=0, final_scores=0.4228915,
final_local_scores=0.55985594, pem=0.0, similarity=20.79693]
73: 31769 United_Nations
    [pem_pos=0, final_scores=0.45439553,
final_local_scores=0.48079318, pem=0.0, similarity=18.43754]
74: 5405 China
    [pem_pos=0, final_scores=0.5098226,
final_local_scores=0.5351676, pem=-0.15669827,
similarity=20.91233]
75: 553623 Great_Hall_of_the_People
    [pem_pos=0, final_scores=0.4624352,
final_local_scores=0.54761726, pem=0.0, similarity=20.431704]
76: 31769 United_Nations
    [pem_pos=0, final_scores=0.47223532,
final_local_scores=0.50028044, pem=0.0, similarity=19.01908]
1: 9209651 Communism
    [pem_pos=0, final_scores=0.049156323,
final_local_scores=0.02589716, pem=-0.39361537,
similarity=7.0030775]
```

Sample: b'700_UK'

UK lowers noise limits for three [1 1 London 1] airports . [2 2 LONDON 2] 1996-08-28 The [3 3 British 3] government on Wednesday lowered the noise limits for [4 4 London 4] 's [5 5 Heathrow 5] , [6 6 Gatwick 6] and [7 7 Stansted 7] airports and announced it would make a bigger effort in detecting and fining violators . The limits , effective from January 1 , 1997 , are reduced as much as possible while still complying with international obligations , a spokesman for the

Department of Transport said . The maximum noise level during the day is trimmed by three decibels to 94 , while the night time level is reduced by two decibels to 87 . " It is a smaller reduction in terms of loudness than was sought by local people . Nevertheless I am satisfied that the overall benefits will be worthwhile , " Lord Goschen , minister for aviation , said in a statement . The ministry said it believed the new limits could be met with existing aircraft . " They can be flown in quieter ways , " a spokesman said . The reduction in noise levels is the same as proposed in a consultation paper which was published in October 1995 . The present noise levels have applied at [8 8 Heathrow 8] , one of the world 's busiest airports , since 1959 and at [9 9 Gatwick 9] since 1968 . The number of monitors will be increased and some will be repositioned to detect noisy planes . --

London Newsroom +44 171 542 7717

```
1: 17867 London
2: 17867 London
3: 31717 United_Kingdom
4: 17867 London
5: 94089 Heathrow_(hamlet)
6: 101359 Gatwick_Airport
7: 65456 London_Stansted_Airport
8: 94089 Heathrow_(hamlet)
9: 101359 Gatwick_Airport
1: 17867 London
    [pem_pos=0, final_scores=0.57506293,
final_local_scores=0.56217134, pem=-0.09086737,
similarity=21.360178]
2: 17867 London
    [pem_pos=0, final_scores=0.6470436,
final_local_scores=0.6407993, pem=-0.09086737,
similarity=23.706593]
3: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5506993,
final_local_scores=0.56192225, pem=-0.48936236,
similarity=23.519823]
4: 17867 London
    [pem_pos=0, final_scores=0.5526509,
final_local_scores=0.5376896, pem=-0.09086737,
similarity=20.629595]
5: 94089 Heathrow_(hamlet)
    [pem_pos=4, final_scores=0.39953905,
final_local_scores=0.43774062, pem=-5.6237597,
similarity=47.735645]
6: 101359 Gatwick_Airport
    [pem_pos=0, final_scores=0.3895591,
final_local_scores=0.37178236, pem=0.0, similarity=15.18444]
7: 65456 London_Stansted_Airport
    [pem_pos=0, final_scores=0.5068398,
final_local_scores=0.5158177, pem=-0.2670293,
similarity=20.934889]
```

```
8: 94089 Heathrow_(hamlet)
    [pem_pos=4, final_scores=0.4215356,
final_local_scores=0.46176857, pem=-5.6237597,
similarity=48.452686]
9: 101359 Gatwick_Airport
    [pem_pos=0, final_scores=0.33310372,
final_local_scores=0.31011346, pem=0.0, similarity=13.344117]
```

Sample: b'701_Tennis'

Tennis - [1 1 Philippoussis 1] beats [2 2 Woodforde 2] in [3 3 U.S. Open 3] . [4 4 NEW YORK 4] 1996-08-27 [5 5 Mark Philip poussis 5] beat fellow [6 6 Australian 6] [7 7 Mark Woodforde 7] 6-7 (6-8) 6-3 6-3 6-3 in a men 's singles first round match at the [8 8 U.S. Open 8] on Tuesday . The two [9 9 Davis Cup 9] team-mates were pitted against each other after last week 's controversial redraw of the men 's singles competition . [10 10 Philippoussis 10] is on course for a third round meeting with world number one [11 11 Pete Sampras 11] of the [12 12 United States 12] , whom he beat at the [13 13 Australian Open 13] last January . -- Sydney Newsroom 61-2 9373-1800

```
1: 261368 Mark_Philippoussis
2: 1918850 Mark_Woodforde
3: 82729 US_Open_(tennis)
4: 8210131 New_York
5: 261368 Mark_Philippoussis
6: 4689264 Australia
7: 1918850 Mark_Woodforde
8: 82729 US_Open_(tennis)
9: 227983 Davis_Cup
10: 261368 Mark_Philippoussis
11: 65961 Pete_Sampras
12: 3434750 United_States
13: 97851 Australian_Open
1: 261368 Mark_Philippoussis
    [pem_pos=0, final_scores=0.7036409,
final_local_scores=0.7053874, pem=0.0, similarity=25.139881]
2: 1918850 Mark_Woodforde
    [pem_pos=0, final_scores=0.75602347,
final_local_scores=0.81293964, pem=0.0, similarity=28.349453]
3: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.5239834,
final_local_scores=0.4935844, pem=-1.2370842,
similarity=25.546711]
4: 8210131 New_York
    [pem_pos=0, final_scores=0.14299189,
final_local_scores=0.216097, pem=-0.2889045,
similarity=12.109586]
5: 261368 Mark_Philippoussis
    [pem_pos=0, final_scores=0.81721383,
final_local_scores=0.82944864, pem=0.0, similarity=28.842115]
6: 4689264 Australia
```

```
[pem_pos=0, final_scores=0.7188806,
final_local_scores=0.8333492, pem=-0.491009,
similarity=31.628698]
7: 1918850 Mark_Woodforde
    [pem_pos=0, final_scores=0.9665744,
final_local_scores=1.0257301, pem=0.0, similarity=34.699547]
8: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.57326907,
final_local_scores=0.5474214, pem=-1.2370842,
similarity=27.153315]
9: 227983 Davis_Cup
    [pem_pos=0, final_scores=0.46809006,
final_local_scores=0.4638486, pem=-0.21573581,
similarity=19.105085]
10: 261368 Mark_Philippoussis
    [pem_pos=0, final_scores=0.61235994,
final_local_scores=0.60567683, pem=0.0, similarity=22.164318]
11: 65961 Pete_Sampras
    [pem_pos=0, final_scores=0.6613117,
final_local_scores=0.65446234, pem=0.0, similarity=23.620174]
12: 3434750 United_States
    [pem_pos=0, final_scores=0.4831962,
final_local_scores=0.56609225, pem=-0.073457584,
similarity=21.38251]
13: 97851 Australian_Open
    [pem_pos=0, final_scores=0.4197023,
final_local_scores=0.39077467, pem=-0.3261775,
similarity=17.525011]
```

Sample: b'702_Rugby'

Rugby union-England given final chance to stay in Five Nations . [1 1 LONDON 1] 1996-08-28 [2 2 England 2] have been given a final chance to remain in the Five Nations ' championship despite striking an exclusive television deal with [3 3 Rupert Murdoch 3] 's Sky television . In a statement on Wednesday , the Four Nations TV Committee said dates had been set for a competition involving [4 4 Scotland 4] , [5 5 Wales 5] , [6 6 Ireland 6] and [7 7 France 7] next year . " Between now and then , discussions will take place in one final attempt to persuade the [8 8 Rugby Football Union 8] to save the Five Nations ' championship in its current form , " the statement said . No further details were immediately available . [9 9 England 9] infuriated their championship colleagues when they decided to sign a 87.5 million pounds sterling (\$ 135.8 million) deal giving Sky television exclusive rights to rugby union matches in [10 10 E ngland 10] . The present contract with the [11 11 British Bro adcasting Corporation 11] was shared between the four home nations while [12 12 France 12] have their own television deal . Last month Five Nations ' committee chairman [13 13 Tom Kiernan 13] said [14 14 England 14] would be thrown out of the competition " unless circumstances change in the near

```
future " .  
1: 17867 London  
2: 691024 England_national_rugby_union_team  
3: 26091 Rupert_Murdoch  
4: 26994 Scotland  
5: 69894 Wales  
6: 147575 Ireland  
7: 5843419 France  
8: 696501 Rugby_Football_Union  
9: 691024 England_national_rugby_union_team  
10: 9316 England  
11: 19344654 BBC  
12: 5843419 France  
13: 14300052 Tom_Kiernan  
14: 9904 England_national_football_team  
1: 17867 London  
    [pem_pos=0, final_scores=0.6183368,  
final_local_scores=0.61225426, pem=-0.09086737,  
similarity=22.854752]  
2: 691024 England_national_rugby_union_team  
    [pem_pos=3, final_scores=0.31593794,  
final_local_scores=0.29197365, pem=-3.7874734,  
similarity=33.39966]  
3: 26091 Rupert_Murdoch  
    [pem_pos=0, final_scores=0.43694746,  
final_local_scores=0.5084214, pem=0.0, similarity=19.262022]  
4: 26994 Scotland  
    [pem_pos=0, final_scores=0.49787438,  
final_local_scores=0.489334, pem=-0.22905664,  
similarity=19.93806]  
5: 69894 Wales  
    [pem_pos=0, final_scores=0.5745433,  
final_local_scores=0.5657915, pem=-0.37222588,  
similarity=22.998281]  
6: 147575 Ireland  
    [pem_pos=0, final_scores=0.33101398,  
final_local_scores=0.29691336, pem=-0.5653214,  
similarity=16.024506]  
7: 5843419 France  
    [pem_pos=0, final_scores=0.5333978,  
final_local_scores=0.5418692, pem=-0.16240834,  
similarity=21.143372]  
8: 696501 Rugby_Football_Union  
    [pem_pos=0, final_scores=0.31758004,  
final_local_scores=0.31930324, pem=0.0, similarity=13.618359]  
9: 691024 England_national_rugby_union_team  
    [pem_pos=3, final_scores=0.40380317,  
final_local_scores=0.3879531, pem=-3.7874734,  
similarity=36.263878]  
10: 9316 England  
    [pem_pos=0, final_scores=0.63434094,  
final_local_scores=0.61522764, pem=-0.19053856,
```

```
similarity=23.485512]
11: 19344654 BBC
    [pem_pos=0, final_scores=0.48309416,
final_local_scores=0.50728244, pem=0.0, similarity=19.228033]
12: 5843419 France
    [pem_pos=0, final_scores=0.5799916,
final_local_scores=0.59276587, pem=-0.16240834,
similarity=22.66223]
13: 14300052 Tom_Kiernan
    [pem_pos=0, final_scores=0.43109185,
final_local_scores=0.536241, pem=0.0, similarity=20.092215]
14: 9904 England_national_football_team
    [pem_pos=1, final_scores=0.3051718,
final_local_scores=0.2963503, pem=-2.776464,
similarity=28.03224]
```

Sample: b'703_Cricket'

Cricket - NZ face tough schedule at home and abroad . [1 1 WELLINGTON 1] 1996-08-28 [2 2 World Cup 2] cricket champions [3 3 Sri Lanka 3] will play two tests and three one-day internationals in a tour of [4 4 New Zealand 4] next March , officials said on Wednesday . [5 5 New Zealand 5] Cricket said the [6 6 Sri Lankans 6] would play tests in [7 7 Hamilton 7] and [8 8 Wellington 8] and one-dayers in [9 9 Auckland 9] , [10 10 Christchurch 10] and [11 11 Dunedin 11] , following hard on the heels of a tour by [12 12 England 12] . [13 13 New Zealand 13] will also line up against [14 14 Sri Lanka 14] and [15 15 Pakistan 15] this November in a one-day champions trophy competition in [16 16 Sharjah 16] . The team will go one to tour [17 17 Pakistan 17] , playing two tests and three one-day internationals .

```
1: 33804 Wellington
2: 7239 Cricket_World_Cup
3: 397246 Sri_Lanka_national_cricket_team
4: 4913064 New_Zealand
5: 4913064 New_Zealand
6: 397246 Sri_Lanka_national_cricket_team
7: 86654 Hamilton,_New_Zealand
8: 33804 Wellington
9: 18660332 Auckland
10: 628782 Christchurch
11: 628345 Dunedin
12: 9316 England
13: 478115 New_Zealand_national_cricket_team
14: 397246 Sri_Lanka_national_cricket_team
15: 759295 Pakistan_national_cricket_team
16: 15674115 Sharjah_(city)
17: 23235 Pakistan
1: 33804 Wellington
    [pem_pos=0, final_scores=0.47158927,
final_local_scores=0.44437838, pem=-0.9265605,
similarity=22.38963]
```

```
2: 7239 Cricket_World_Cup
    [pem_pos=4, final_scores=0.478777,
final_local_scores=0.49595726, pem=-3.1731894,
similarity=36.146362]
3: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.68488973,
final_local_scores=0.71144885, pem=-2.2667584,
similarity=37.64775]
4: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.5439343,
final_local_scores=0.5409356, pem=-0.23795462,
similarity=21.526344]
5: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.5025373,
final_local_scores=0.49571562, pem=-0.23795462,
similarity=20.17689]
6: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=0, final_scores=0.6779786,
final_local_scores=0.70389956, pem=-0.17988552,
similarity=26.073725]
7: 86654 Hamilton,_New_Zealand
    [pem_pos=1, final_scores=0.43912297,
final_local_scores=0.41654307, pem=-2.40051,
similarity=29.574535]
8: 33804 Wellington
    [pem_pos=0, final_scores=0.6133589,
final_local_scores=0.5992402, pem=-0.9265605,
similarity=27.011015]
9: 18660332 Auckland
    [pem_pos=0, final_scores=0.6065554,
final_local_scores=0.58690935, pem=-0.297032,
similarity=23.219566]
10: 628782 Christchurch
    [pem_pos=0, final_scores=0.5117219,
final_local_scores=0.48915046, pem=-0.16020472,
similarity=19.558157]
11: 628345 Dunedin
    [pem_pos=0, final_scores=0.47853965,
final_local_scores=0.4584155, pem=-0.059943043,
similarity=18.095726]
12: 9316 England
    [pem_pos=0, final_scores=0.452493,
final_local_scores=0.45160294, pem=-0.19053856,
similarity=18.602623]
13: 478115 New_Zealand_national_cricket_team
    [pem_pos=1, final_scores=0.4474337,
final_local_scores=0.43334866, pem=-3.142959,
similarity=34.1136]
14: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.80213445,
final_local_scores=0.839521, pem=-2.2667584,
similarity=41.469677]
```

```
15: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7268562,
final_local_scores=0.7522351, pem=-2.7556365,
similarity=41.523487]
16: 15674115 Sharjah_(city)
    [pem_pos=0, final_scores=0.17137556,
final_local_scores=0.1815702, pem=-0.425711,
similarity=11.823213]
17: 23235 Pakistan
    [pem_pos=0, final_scores=0.5117494,
final_local_scores=0.55048406, pem=-0.21944217,
similarity=21.710615]
```

Sample: b'704_Soccer'

Soccer - [1 1 Burundi 1] disqualification from [2 2 African Cup 2] confirmed . [3 3 CAIRO 3] 1996-08-28 The [4 4 African Football Confederation 4] ([5 5 CAF 5]) on Wednesday formally confirmed [6 6 Burundi 6] 's disqualification from the [7 7 African Nations Cup 7] following the team 's inability to travel for a qualifier against [8 8 Central African Republic 8] . The [9 9 Burundi 9] team were unable to leave their troubled country for a preliminary round first leg match in [10 10 Bangui 10] earlier this month because of an air ban imposed in a recent set of internationally-sponsored sanctions . The [11 11 Central African Republic 11] qualified on a walkover to play in group four with [12 12 Guinea 12] , [13 13 Sierra Leone 13] and [14 14 Tunisia 14] . " After examining the dossier of the Burundi-Central Africa match , we decided ... to disqualify the national team of [15 15 Burundi 15] from the 21st African Cup of Nations ... as a result of the absence of this team from the match , " [16 16 CAF 16] said in a statement .

```
1: 1212797 Burundi_national_football_team
2: 250680 Africa_Cup_of_Nations
3: 6293 Cairo
4: 715265 Confederation_of_African_Football
5: 715265 Confederation_of_African_Football
6: 1212797 Burundi_national_football_team
7: 250680 Africa_Cup_of_Nations
8: 5478 Central_African_Republic
9: 1212797 Burundi_national_football_team
10: 56617 Bangui
11: 5478 Central_African_Republic
12: 1076998 Guinea_national_football_team
13: 1145024 Sierra_Leone_national_football_team
14: 1029953 Tunisia_national_football_team
15: 1212797 Burundi_national_football_team
16: 715265 Confederation_of_African_Football
1: 1212797 Burundi_national_football_team
    [pem_pos=1, final_scores=0.64165026,
final_local_scores=0.6455018, pem=-2.2049255,
similarity=35.343502]
```

```
2: 250680 Africa_Cup_of_Nations
    [pem_pos=0, final_scores=0.4327938,
final_local_scores=0.39080983, pem=0.0, similarity=15.752258]
3: 6293 Cairo
    [pem_pos=0, final_scores=0.46211714,
final_local_scores=0.5316298, pem=-0.15211283,
similarity=20.78182]
4: 715265 Confederation_of_African_Football
    [pem_pos=0, final_scores=0.46891636,
final_local_scores=0.4390154, pem=0.0, similarity=17.190807]
5: 715265 Confederation_of_African_Football
    [pem_pos=0, final_scores=0.37346148,
final_local_scores=0.33474547, pem=-0.26986858,
similarity=15.546773]
6: 1212797 Burundi_national_football_team
    [pem_pos=1, final_scores=0.574038,
final_local_scores=0.5716457, pem=-2.2049255,
similarity=33.13949]
7: 250680 Africa_Cup_of_Nations
    [pem_pos=0, final_scores=0.53661543,
final_local_scores=0.50129163, pem=0.0, similarity=19.049255]
8: 5478 Central_African_Republic
    [pem_pos=0, final_scores=0.48023725,
final_local_scores=0.53338, pem=-0.30754697,
similarity=21.679323]
9: 1212797 Burundi_national_football_team
    [pem_pos=1, final_scores=0.5235166,
final_local_scores=0.5164587, pem=-2.2049255,
similarity=31.492598]
10: 56617 Bangui
    [pem_pos=0, final_scores=0.3969695,
final_local_scores=0.45206386, pem=0.0, similarity=17.5802]
11: 5478 Central_African_Republic
    [pem_pos=0, final_scores=0.47092903,
final_local_scores=0.52321213, pem=-0.30754697,
similarity=21.375896]
12: 1076998 Guinea_national_football_team
    [pem_pos=1, final_scores=0.43096137,
final_local_scores=0.41486955, pem=-2.1209195,
similarity=28.004137]
13: 1145024 Sierra_Leone_national_football_team
    [pem_pos=1, final_scores=0.5779379,
final_local_scores=0.60244375, pem=-2.135507,
similarity=33.681057]
14: 1029953 Tunisia_national_football_team
    [pem_pos=1, final_scores=0.43082148,
final_local_scores=0.41750365, pem=-2.4684548,
similarity=29.972694]
15: 1212797 Burundi_national_football_team
    [pem_pos=1, final_scores=0.6333587,
final_local_scores=0.6364445, pem=-2.2049255,
similarity=35.073215]
```

```
16: 715265 Confederation_of_African_Football
    [pem_pos=0, final_scores=0.38526875,
final_local_scores=0.34764308, pem=-0.26986858,
similarity=15.931663]

Sample: b'705_Cricket'
Cricket - [1 1 India 1] 226-5 in 50 overs v [2 2 Sri Lanka 2]
. [3 3 COLOMBO 3] 1996-08-28 [4 4 India 4] scored 226 for
five wickets in their 50 overs against [5 5 Sri Lanka 5] in
the second day-night limited overs match of the [6 6 Singer Wo
rld Series 6] tournament on Wednesday .
1: 407754 India_national_cricket_team
2: 397246 Sri_Lanka_national_cricket_team
3: 56636 Colombo
4: 407754 India_national_cricket_team
5: 397246 Sri_Lanka_national_cricket_team
6: 28357832 Singer_World_Series
1: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.69567806,
final_local_scores=0.65638757, pem=-3.5152702,
similarity=42.794212]
2: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.70430964,
final_local_scores=0.66331404, pem=-2.2667584,
similarity=36.21131]
3: 56636 Colombo
    [pem_pos=0, final_scores=0.50774306,
final_local_scores=0.56245965, pem=-0.19393882,
similarity=21.9293]
4: 407754 India_national_cricket_team
    [pem_pos=1, final_scores=0.7782653,
final_local_scores=0.7466016, pem=-3.5152702,
similarity=45.48638]
5: 397246 Sri_Lanka_national_cricket_team
    [pem_pos=1, final_scores=0.8578927,
final_local_scores=0.83108014, pem=-2.2667584,
similarity=41.217785]
6: 28357832 Singer_World_Series
    [pem_pos=0, final_scores=0.64062387,
final_local_scores=0.7081184, pem=0.0, similarity=25.221378]
```

```
Sample: b'706_Canada'
[1 1 Canada 1] fast-tracks [2 2 Chinese 2] asylum-seekers -
report . [3 3 VANCOUVER 3] , [4 4 British Columbia 4]
1996-08-28 [5 5 Canada 5] is fast-tracking immigration
applications from [6 6 Chinese 6] dissidents in [7 7 Hong Kong
7] before the [8 8 British 8] colony reverts to [9 9 China 9]
's control next year , the [10 10 Vancouver Sun 10] reported
on Wednesday . The applications are being " fast-tracked in
the sense that we are processing them and the ones who have
been referred to us have been interviewed , " the newspaper
quoted Garrett Lambert , [11 11 Canada 11] 's high
```

commissioner in [12 12 Hong Kong 12] , as saying . " A small number already have preliminary indications as to what the disposition of their cases are and so I suppose in that sense , I guess we have given them some preferential treatment , " Lambert said . He declined to say how many people were being considered for asylum . [13 13 Canada 13] 's ministry of foreign affairs in [14 14 Ottawa 14] had no immediate comment on the report . About 80 [15 15 Chinese 15] dissidents are believed to be living in exile in [16 16 Hong Kong 16] . Their fate after the territory reverts to [17 17 Chinese 17] rule is unclear . [18 18 Britain 18] hands [19 19 Hong Kong 19] back to [20 20 China 20] at midnight on June 30 , 1997 , after 150 years of colonial rule . [21 21 Canada 21] 's Minister of Foreign Affairs [22 22 Lloyd Axworthy 22] said after meeting [23 23 Hong Kong 23] Gov . [24 24 Chris Patten 24] last month that [25 25 Canada 25] may grant asylum to dissidents who have fled to [26 26 Hong Kong 26] from [27 27 China 27] . [28 28 Chinese 28] officials have said such dissidents may not become [29 29 Hong Kong 29] permanent residents since they entered the territory illegally but have also said their status would be decided by the post-1997 local [30 30 Hong Kong 30] administration .

1: 5042916 Canada
2: 5405 China
3: 32706 Vancouver
4: 3392 British_Columbia
5: 5042916 Canada
6: 5405 China
7: 13404 Hong_Kong
8: 31717 United_Kingdom
9: 5405 China
10: 233978 The_Vancouver_Sun
11: 5042916 Canada
12: 13404 Hong_Kong
13: 5042916 Canada
14: 22219 Ottawa
15: 5405 China
16: 13404 Hong_Kong
17: 5405 China
18: 31717 United_Kingdom
19: 13404 Hong_Kong
20: 5405 China
21: 5042916 Canada
22: 427533 Lloyd_Axworthy
23: 13404 Hong_Kong
24: 201835 Chris_Patten
25: 5042916 Canada
26: 13404 Hong_Kong
27: 5405 China
28: 5405 China
29: 13404 Hong_Kong
30: 13404 Hong_Kong

```
1: 5042916 Canada
    [pem_pos=0, final_scores=0.6724569,
final_local_scores=0.6458832, pem=-0.11662793,
similarity=23.998396]
2: 5405 China
    [pem_pos=2, final_scores=0.72412366,
final_local_scores=0.7088755, pem=-1.4196125,
similarity=32.964046]
3: 32706 Vancouver
    [pem_pos=0, final_scores=0.4297502,
final_local_scores=0.4069515, pem=-0.12074051,
similarity=16.890564]
4: 3392 British_Columbia
    [pem_pos=0, final_scores=0.44114643,
final_local_scores=0.42242813, pem=-0.028628228,
similarity=16.851496]
5: 5042916 Canada
    [pem_pos=0, final_scores=0.78792983,
final_local_scores=0.7720198, pem=-0.11662793,
similarity=27.762564]
6: 5405 China
    [pem_pos=2, final_scores=0.7772191,
final_local_scores=0.76687425, pem=-1.4196125,
similarity=34.694843]
7: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.51552933,
final_local_scores=0.46335638, pem=-0.22402541,
similarity=19.135477]
8: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5374451,
final_local_scores=0.509919, pem=-0.48936236,
similarity=21.967941]
9: 5405 China
    [pem_pos=0, final_scores=0.55890626,
final_local_scores=0.5284006, pem=-0.15669827,
similarity=20.71039]
10: 233978 The_Vancouver_Sun
    [pem_pos=0, final_scores=0.44084334,
final_local_scores=0.52574044, pem=0.0, similarity=19.778856]
11: 5042916 Canada
    [pem_pos=0, final_scores=0.7271536,
final_local_scores=0.705631, pem=-0.11662793,
similarity=25.781391]
12: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5393283,
final_local_scores=0.48935312, pem=-0.22402541,
similarity=19.911272]
13: 5042916 Canada
    [pem_pos=0, final_scores=0.6980503,
final_local_scores=0.67384005, pem=-0.11662793,
similarity=24.832684]
14: 22219 Ottawa
```

```
[pem_pos=0, final_scores=0.4482414,
final_local_scores=0.43219835, pem=-0.34965137,
similarity=18.888832]
15: 5405 China
    [pem_pos=2, final_scores=0.58929116,
final_local_scores=0.56159145, pem=-1.4196125,
similarity=28.568794]
16: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.51385826,
final_local_scores=0.46153098, pem=-0.22402541,
similarity=19.081003]
17: 5405 China
    [pem_pos=2, final_scores=0.84560496,
final_local_scores=0.84157544, pem=-1.4196125,
similarity=36.924076]
18: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.56262296,
final_local_scores=0.537422, pem=-0.61114657,
similarity=23.450968]
19: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.54171497,
final_local_scores=0.4919603, pem=-0.22402541,
similarity=19.989075]
20: 5405 China
    [pem_pos=0, final_scores=0.49537224,
final_local_scores=0.45899934, pem=-0.15669827,
similarity=18.639318]
21: 5042916 Canada
    [pem_pos=0, final_scores=0.7914943,
final_local_scores=0.7759135, pem=-0.11662793,
similarity=27.87876]
22: 427533 Lloyd_Axworthy
    [pem_pos=0, final_scores=0.64710695,
final_local_scores=0.7789124, pem=0.0, similarity=27.334013]
23: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.56146246,
final_local_scores=0.5135314, pem=-0.22402541,
similarity=20.6328]
24: 201835 Chris_Patten
    [pem_pos=0, final_scores=0.56214917,
final_local_scores=0.6508789, pem=0.0, similarity=23.513237]
25: 5042916 Canada
    [pem_pos=0, final_scores=0.7283576,
final_local_scores=0.70694625, pem=-0.11662793,
similarity=25.82064]
26: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5438586,
final_local_scores=0.4943018, pem=-0.22402541,
similarity=20.05895]
27: 5405 China
    [pem_pos=0, final_scores=0.56836385,
final_local_scores=0.5387316, pem=-0.15669827,
```

```
similarity=21.018688]
28: 5405 China
    [pem_pos=2, final_scores=0.73841494,
final_local_scores=0.72448665, pem=-1.4196125,
similarity=33.429913]
29: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.46187705,
final_local_scores=0.40474945, pem=-0.22402541,
similarity=17.38653]
30: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.51890045,
final_local_scores=0.4670388, pem=-0.22402541,
similarity=19.245369]
```

Sample: b'707_Ivorian'

[1 1 Ivorian 1] journalist held , asked to reveal source . [2 2 ABIDJAN 2] 1996-08-28 An [3 3 Ivorian 3] journalist spent a third day in custody on Wednesday and investigators were demanding that he reveal the source of an official document published in his newspaper , colleagues said . Raphael Lapke , publication director of [4 4 Ivorian 4] newspaper Le Populaire , was taken in for questioning on Monday over an article about the public prosecutor and has been detained since then . Colleagues said he had been charged with theft of administrative documents . " He is being asked for the source of his information and who gave him this confidential document , " one colleague told [1 Reuters 1] . Three journalists from the [5 5 Ivorian 5] opposition daily La Voie are serving two-year prison terms for insulting President [6 6 Henri Konan Bedie 6] . A court sentenced two in December and the third in January . La Voie published an article suggesting the presence of Bedie had brought local team [7 7 A SEC 7] bad luck during their defeat by [8 8 Orlando Pirates 8] of [9 9 South Africa 9] in the final of the African Champions Cup in December . The [10 10 United States 10] embassy in [11 11 Abidjan 11] and international press organisations denounced the sentences as excessive . Last year , Bedie pardoned four journalists jailed for the same or similar offences . They included one of the three La Voie journalists . Bedie pardoned two other journalists jailed for incitement to disturb public order .

```
1: 57744 Ivory_Coast
2: 21371767 Abidjan
3: 57744 Ivory_Coast
4: 57744 Ivory_Coast
5: 57744 Ivory_Coast
6: 418112 Henri_Konan_Bédié
7: 2870067 ASEC_Mimosas
8: 334800 Orlando_Pirates
9: 17416221 South_Africa
10: 3434750 United_States
11: 21371767 Abidjan
```

```
1: 57744 Ivory_Coast
    [pem_pos=0, final_scores=0.4580776,
final_local_scores=0.4110294, pem=0.0, similarity=16.35565]
2: 21371767 Abidjan
    [pem_pos=0, final_scores=0.42517012,
final_local_scores=0.38733816, pem=0.0, similarity=15.648657]
3: 57744 Ivory_Coast
    [pem_pos=0, final_scores=0.5184124,
final_local_scores=0.47693604, pem=0.0, similarity=18.322435]
4: 57744 Ivory_Coast
    [pem_pos=0, final_scores=0.55667883,
final_local_scores=0.51873624, pem=0.0, similarity=19.569838]
5: 57744 Ivory_Coast
    [pem_pos=0, final_scores=0.55197626,
final_local_scores=0.51359946, pem=0.0, similarity=19.416546]
6: 418112 Henri_Konan_Bédié
    [pem_pos=0, final_scores=0.6500647,
final_local_scores=0.7762126, pem=0.0, similarity=27.253445]
7: 2870067 ASEC_Mimosas
    [pem_pos=1, final_scores=0.41988444,
final_local_scores=0.4802825, pem=-0.24976285,
similarity=19.78055]
8: 334800 Orlando_Pirates
    [pem_pos=0, final_scores=0.29826105,
final_local_scores=0.354416, pem=0.0, similarity=14.666194]
9: 17416221 South_Africa
    [pem_pos=0, final_scores=0.47265905,
final_local_scores=0.48821825, pem=-0.26251382,
similarity=20.08671]
10: 3434750 United_States
    [pem_pos=0, final_scores=0.4391021,
final_local_scores=0.47136372, pem=-0.073457584,
similarity=18.55562]
11: 21371767 Abidjan
    [pem_pos=0, final_scores=0.3921641,
final_local_scores=0.35128412, pem=0.0, similarity=14.572732]
1: 18998750 Reuters
    [pem_pos=0, final_scores=0.2402934,
final_local_scores=0.30792716, pem=0.0, similarity=13.278873]
```

Sample: b'708_Village'
Village attack kills 38 in eastern [1 1 Sierra Leone 1] . [2
2 FREETOWN 2] 1996-08-28 [3 3 Sierra Leonean 3] rebels killed
31 villagers and seven soldiers in an attack on the eastern
village of [4 4 Foindu 4] , Eastern Region Brigade Commander
Major Fallah Sewa said on Wednesday . Sewa said the rebels
overran [5 5 Foindu 5] despite the presence of government
troops in the village on the highway between Mano Junction
and the diamond town of Tongo Field . An army spokesman in [6
6 Freetown 6] said Monday night 's attack was the third on a
military post in the past week . Rebels of the [7 7 Revolutionary
United Front 7] agreed a ceasefire in April . Continuing

attacks are generally ascribed to renegade soldiers or uncontrolled bands of rebels and refugees displaced by the fighting starting to return to their homes . Peace talks in [8 8 Ivory Coast 8] began in February . Diplomats say they are deadlocked over the [9 9 RUF 9] 's insistence that foreign troops helping the government army should leave , and that they should have some say in the allocation of budget spending .

1: 27308 Sierra_Leone
2: 56661 Freetown
3: 27308 Sierra_Leone
4: 14659758 Foindu
5: 14659758 Foindu
6: 56661 Freetown
7: 298010 Revolutionary_United_Front
8: 57744 Ivory_Coast
9: 298010 Revolutionary_United_Front
1: 27308 Sierra_Leone
[pem_pos=0, final_scores=0.48889273,
final_local_scores=0.46856833, pem=-0.29412922,
similarity=19.672247]
2: 56661 Freetown
[pem_pos=0, final_scores=0.42084426,
final_local_scores=0.4129157, pem=-0.08093871,
similarity=16.852098]
3: 27308 Sierra_Leone
[pem_pos=0, final_scores=0.40409112,
final_local_scores=0.37593544, pem=0.0, similarity=15.308376]
4: 14659758 Foindu
[pem_pos=0, final_scores=0.48893327,
final_local_scores=0.6415377, pem=0.0, similarity=23.234478]
5: 14659758 Foindu
[pem_pos=0, final_scores=0.52753603,
final_local_scores=0.6837054, pem=0.0, similarity=24.492846]
6: 56661 Freetown
[pem_pos=0, final_scores=0.3339482,
final_local_scores=0.31799495, pem=-0.08093871,
similarity=14.019474]
7: 298010 Revolutionary_United_Front
[pem_pos=0, final_scores=0.48759776,
final_local_scores=0.49897003, pem=0.0, similarity=18.979975]
8: 57744 Ivory_Coast
[pem_pos=0, final_scores=0.44536847,
final_local_scores=0.4691242, pem=0.0, similarity=18.089315]
9: 298010 Revolutionary_United_Front
[pem_pos=0, final_scores=0.36781543,
final_local_scores=0.36812603, pem=-0.19246598,
similarity=16.121988]

Sample: b'709_Aid'

Aid agency says [1 1 Sudan 1] missionaries released . [2 2 NA IROBI 2] 1996-08-28 An aid agency said six [3 3 Roman Catholic

3] missionaries , including three [4 4 Australian 4] nuns , were freed by rebels in southern [5 5 Sudan 5] on Wednesday after being held for nearly two weeks . But [6 6 Catholic 6] church officials said they had no confirmation of the report and would have to wait until Thursday to be sure .

1: 27421 Sudan
2: 21482 Nairobi
3: 606848 Catholic_Church
4: 4689264 Australia
5: 27421 Sudan
6: 606848 Catholic_Church
1: 27421 Sudan
[pem_pos=0, final_scores=0.5741886, final_local_scores=0.5634216, pem=-0.2848764, similarity=22.45254]
2: 21482 Nairobi
[pem_pos=0, final_scores=0.42811412, final_local_scores=0.44043767, pem=0.0, similarity=17.233252]
3: 606848 Catholic_Church
[pem_pos=0, final_scores=0.38354266, final_local_scores=0.40131384, pem=0.0, similarity=16.065718]
4: 4689264 Australia
[pem_pos=0, final_scores=0.5214724, final_local_scores=0.54918414, pem=-0.491009, similarity=23.148645]
5: 27421 Sudan
[pem_pos=0, final_scores=0.5517433, final_local_scores=0.5389036, pem=-0.2848764, similarity=21.720873]
6: 606848 Catholic_Church
[pem_pos=3, final_scores=0.14918189, final_local_scores=0.14531031, pem=-0.9041281, similarity=13.342852]

Sample: b'710_Zambia'

[1 1 Zambia 1] 's [2 2 Chiluba 2] shuffles cabinet to fill vacancy . [1 LUSAKA 1] 1996-08-28 Zambian President [3 3 Fred erick Chiluba 3] shuffled his cabinet on Wednesday to fill a vacancy left after the sacking of Legal Affairs Minister Remmy Mushota . Mushota was fired a month ago after a government tribunal found he tried to withdraw cash from state coffers without authority . The president 's office said in a statement that Lands Minister Luminzu Shimaponda had been appointed Legal Affairs Minister , while Deputy Foreign Minister Peter Machungwa would take over from Shimaponda .

1: 34415 Zambia
2: 285508 Frederick_Chiluba
3: 285508 Frederick_Chiluba
1: 34415 Zambia
[pem_pos=0, final_scores=0.48260826, final_local_scores=0.52465945, pem=-0.32073748,

```
similarity=21.490818]
2: 285508 Frederick_Chiluba
    [pem_pos=0, final_scores=0.5214469,
final_local_scores=0.5057968, pem=0.0, similarity=19.183699]
3: 285508 Frederick_Chiluba
    [pem_pos=0, final_scores=0.88035434,
final_local_scores=0.89784867, pem=0.0, similarity=30.883308]
1: 23149963 Lusaka
    [pem_pos=0, final_scores=0.07248007,
final_local_scores=0.055835083, pem=0.0, similarity=5.7559433]
```

Sample: b'711_Guinea'

Guinea launches war on fictitious civil servants . [1 1 CONA KRY 1] 1996-08-28 Guinea launched a drive on Wednesday to rid the civil service payroll of fictitious workers as part of new prime minister [2 2 Sidia Toure 2] 's campaign to cut government spending . Deputy Minister for Finance Ousmane Kaba said teams of inspectors would check government offices in the capital and the provinces to root out civil servants who drew salaries but had left their jobs , were dead , or had never existed . Some 50 million Guinean francs (\$ 50,000) has been pumped into the exercise to deter the inspectors from taking bribes . Kaba told reporters the annual wage bill of 171 billion Guinean francs represented 50 percent of current state expenditure , " whereas the acceptable proportion in countries similar to ours is one third " . Toure , who took office last month , has said he plans to cut public service spending by 30 percent by the end of the year as part of measures to revive the economy . Guinea is rich in minerals and has a vast potential for hydroelectric power generation but it faces stiff competition from its West African neighbours for foreign investment . President [3 3 Lansana Conte 3] appointed Toure , a former senior civil servant in [4 4 Ivory Coast 4] , last month to clean up the administration and reform the economy following February 's bloody army revolt . (\$ =1,000 Guinean francs)

```
1: 56638 Conakry
2: 3770586 Sidya_Touré
3: 392412 Lansana_Conté
4: 57744 Ivory_Coast
1: 56638 Conakry
    [pem_pos=0, final_scores=0.38492054,
final_local_scores=0.39907306, pem=0.0, similarity=15.99885]
2: 3770586 Sidya_Touré
    [pem_pos=0, final_scores=0.5150257,
final_local_scores=0.5836245, pem=0.0, similarity=21.506231]
3: 392412 Lansana_Conté
    [pem_pos=0, final_scores=0.5079047,
final_local_scores=0.54673463, pem=0.0, similarity=20.405365]
4: 57744 Ivory_Coast
    [pem_pos=0, final_scores=0.41678357,
final_local_scores=0.47116435, pem=0.0, similarity=18.150196]
```

Sample: b'712_New'

New Liberia Council chief to be installed Tuesday . [1 1 MON ROVIA 1] 1996-08-28 [2 2 Ruth Perry 2] , the woman with the task of uniting [3 3 Liberia 3] 's squabbling factions around the latest peace plan , will be formally installed as head of the ruling State Council next Tuesday , a Council statement said . [4 4 Perry 4] , a [5 5 Liberian 5] Senate member during the 1980s , returned to [6 6 Monrovia 6] on August 22 after West African leaders nominated her for the job under a peace deal signed in [7 7 Nigeria 7] 's capital [8 8 Abuja 8] five days earlier . The formal inauguration had been due to take place this week but was put back . There was no official explanation but politicians said faction leaders and State Council vice-chairmen [9 9 Charles Taylor 9] and Alhadji Kromah were unable to attend because they were travelling . [1 0 10 Liberia 10] 's civil war , launched by Taylor in 1989 , has killed well over 150,000 people . Faction fighting and an orgy of looting in the capital [11 11 Monrovia 11] in April and May killed hundreds of people . Over a dozen peace deals have collapsed . The latest sets a timetable for disarmament by the end of January and elections by May 30 . West African leaders have threatened individual sanctions against faction leaders to ensure compliance . Freed [12 12 American 12] slaves founded [13 13 Liberia 13] in 1847 .

1: 285213 Monrovia

2: 1981141 Ruth_Perry

3: 17791 Liberia

4: 1981141 Ruth_Perry

5: 17791 Liberia

6: 285213 Monrovia

7: 21383 Nigeria

8: 56823 Abuja

9: 241458 Charles_Taylor_(Liberian_politician)

10: 17791 Liberia

11: 285213 Monrovia

12: 3434750 United_States

13: 17791 Liberia

1: 285213 Monrovia

[pem_pos=0, final_scores=0.46136647, final_local_scores=0.41921633, pem=-0.2167236, similarity=17.778542]

2: 1981141 Ruth_Perry

[pem_pos=0, final_scores=0.4857626, final_local_scores=0.52901816, pem=0.0, similarity=19.87667]

3: 17791 Liberia

[pem_pos=0, final_scores=0.55838805, final_local_scores=0.51242846, pem=-0.26536247, similarity=20.824682]

4: 1981141 Ruth_Perry

[pem_pos=0, final_scores=0.3493225, final_local_scores=0.37997812, pem=0.0, similarity=15.429018]

```
5: 17791 Liberia
    [pem_pos=0, final_scores=0.55226165,
final_local_scores=0.50573623, pem=0.0, similarity=19.18189]
6: 285213 Monrovia
    [pem_pos=0, final_scores=0.4368953,
final_local_scores=0.39248526, pem=-0.2167236,
similarity=16.980833]
7: 21383 Nigeria
    [pem_pos=0, final_scores=0.6306492,
final_local_scores=0.6257594, pem=-0.31455088,
similarity=24.4742]
8: 56823 Abuja
    [pem_pos=0, final_scores=0.3950506,
final_local_scores=0.42265618, pem=0.0, similarity=16.702616]
9: 241458 Charles_Taylor_(Liberian_politician)
    [pem_pos=0, final_scores=0.47587532,
final_local_scores=0.5057819, pem=-0.6882932,
similarity=22.926298]
10: 17791 Liberia
    [pem_pos=0, final_scores=0.50400764,
final_local_scores=0.45302606, pem=-0.26536247,
similarity=19.051996]
11: 285213 Monrovia
    [pem_pos=0, final_scores=0.41863078,
final_local_scores=0.3725341, pem=-0.2167236,
similarity=16.38545]
12: 3434750 United_States
    [pem_pos=0, final_scores=0.43146175,
final_local_scores=0.44759744, pem=-0.12994486,
similarity=18.153574]
13: 17791 Liberia
    [pem_pos=0, final_scores=0.47883385,
final_local_scores=0.4255275, pem=-0.26536247,
similarity=18.231384]
```

Sample: b'713_Nigeria'

[1 1 Nigeria 1] rights group says four academics arrested . [2 2 LAGOS 2] 1996-08-28 A [3 3 Nigerian 3] human rights group said on Wednesday that four members of a recently banned university union had been arrested . " The Constitutional Rights Project (CRP) believes that William Istafanus , Elisha Shamay , O.K. Likkason and Jerome Egurugbe were arrested because of their role in the ongoing ASUU ([4 4 Academic Staff Union of Universities 4]) strike , " the group said in a statement . The CRP said the four were arrested on Monday night at the northeastern Tafawa Balewa University . The main academic union , ASUU , along with two smaller university unions , was banned by [5 5 Nigeria 5] 's military government last week , because of a four-month strike by teachers for better working conditions . [6 6 Nigeria 6] is under fire from many [7 7 Western 7] countries for human rights abuses and lack of democracy . Dozens of people

opposed to the government are in detention . [8 8 Commonwealth
8] foreign ministers are to meet in [9 9 London 9] on
Wednesday to discuss what action to take , after a visit to [1
0 10 Nigeria 10] was called off when the government imposed
strict rules on whom the mission would be allowed to see . [11
11 Nigeria 11] was suspended from the club of [12 12 Britain 1
2] and its former colonies in November after the hanging of
nine minority rights activists for murder in spite of
international pleas for clemency .

1: 21383 Nigeria
2: 85232 Lagos
3: 21383 Nigeria
4: 26327468 Academic_Staff_Union_of_Universities
5: 21383 Nigeria
6: 21383 Nigeria
7: 21208200 Western_world
8: 21175158 Commonwealth_of_Nations
9: 17867 London
10: 21383 Nigeria
11: 21383 Nigeria
12: 31717 United_Kingdom
1: 21383 Nigeria
[pem_pos=0, final_scores=0.60762584,
final_local_scores=0.55734855, pem=-0.31455088,
similarity=22.432682]
2: 85232 Lagos
[pem_pos=0, final_scores=0.4578938,
final_local_scores=0.43866503, pem=-0.20534219,
similarity=18.297035]
3: 21383 Nigeria
[pem_pos=0, final_scores=0.6308712,
final_local_scores=0.58274055, pem=-0.075027294,
similarity=21.887863]
4: 26327468 Academic_Staff_Union_of_Universities
[pem_pos=0, final_scores=0.455186,
final_local_scores=0.5941211, pem=0.0, similarity=21.819471]
5: 21383 Nigeria
[pem_pos=0, final_scores=0.70656615,
final_local_scores=0.66542584, pem=-0.31455088,
similarity=25.657925]
6: 21383 Nigeria
[pem_pos=0, final_scores=0.5584905,
final_local_scores=0.5036757, pem=-0.31455088,
similarity=20.830976]
7: 21208200 Western_world
[pem_pos=0, final_scores=0.45591825,
final_local_scores=0.50916207, pem=-1.5917408,
similarity=27.940258]
8: 21175158 Commonwealth_of_Nations
[pem_pos=0, final_scores=0.45087153,
final_local_scores=0.4727928, pem=-0.69133765,
similarity=21.958395]

```
9: 17867 London
    [pem_pos=0, final_scores=0.4981374,
final_local_scores=0.5179002, pem=-0.09086737,
similarity=20.03904]
10: 21383 Nigeria
    [pem_pos=0, final_scores=0.6598787,
final_local_scores=0.61442685, pem=-0.31455088,
similarity=24.136013]
11: 21383 Nigeria
    [pem_pos=0, final_scores=0.56896716,
final_local_scores=0.51511985, pem=-0.31455088,
similarity=21.172493]
12: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5500794,
final_local_scores=0.5476826, pem=-0.61114657,
similarity=23.757164]
```

Sample: b'714_Dutroux'

[1 1 Dutroux 1] suspected in murder of [2 2 Slovak 2]
woman . Peter Laca [3 3 BRATISLAVA 3] 1996-08-28 [4 4 Marc Du
troux 4] , the chief accused in a [5 5 Belgian 5] child
murder and sex abuse scandal , is suspected of murdering a
young [6 6 Slovak 6] woman , the [7 7 Slovak 7] office of
Interpol said on Wednesday . Rudolf Gajdos , head of [8 8 Slo
vak 8] Interpol , told a news conference [9 9 Dutroux 9] was
also believed to have planned the kidnapping of at least one [
10 10 Slovak 10] woman . " One of the police versions in the
case of the murder of young gypsy woman in [11 11 Topolcany 11]
] , western [12 12 Slovakia 12] , this July , is a suspicion
that [13 13 Mark Dutroux 13] could have been involved in the
murder , " Gajdos said without elaborating on the age of the
victim and on the other versions . [14 14 Slovak 14] police ,
Interpol , and [15 15 Belgian 15] police have been following
leads on [16 16 Dutroux 16] 's activities in [17 17 Slovakia 1
7] and the neighbouring [18 18 Czech Republic 18] where he is
known to have made frequent visits . Gajdos said the police
sketch of the suspected murderer was " 60 percent identical
with [19 19 Dutroux 19] 's portrait " , and that [20 20 Dutro
ux 20] was known to have been in [21 21 Topolcany 21] around
the time of the woman 's murder . " [22 22 Topolcany 22] and
the area around this town were reported to have been the most
visited places by [23 23 Dutroux 23] and his accomplices in [2
4 24 Slovakia 24] , " Gajdos said . [25 25 Dutroux 25] , a
convicted child rapist and unemployed father-of-three , led
police 11 days ago to the bodies of eight-year-olds Julie
Lejeune and Melissa Russo in the garden of another of the six
houses he owns around the southern [26 26 Belgian 26] city of
[27 27 Charleroi 27] . " The [28 28 Belgian 28] police also
informed us that [29 29 Dutroux 29] , together with one other
man , had (also) planned the kidnapping of at least one [30
30 Slovak 30] woman , " Gajdos said . " The plan apparently
failed due to difficulties in crossing the border , " he

added , but did not elaborate . The [31 31 Slovak 31] police are also investigating visits by about 10 [32 32 Slovak 32] women , aged 17 to 22 , to [33 33 Belgium 33] , at the invitation of [34 34 Dutroux 34] . The women said they went to [35 35 Belgium 35] voluntarily and police suspect they were used to act in pornographic films , Gajdos said earlier this week . But he added they had difficulty remembering what happened during their visits to [36 36 Belgium 36] , perhaps because of drugs , and were unsure whether they were filmed for pornography . [37 37 Dutroux 37] , 39 , who was charged last week with the abduction and illegal imprisonment of two girls aged 14 and 12 , is also suspected in the disappearance of Belgians An Marchal , 19 , and Eefje Lambrecks , 17 , who went missing a year ago .

1: 154694 Marc_Dutroux
2: 26830 Slovakia
3: 18933194 Bratislava
4: 154694 Marc_Dutroux
5: 3343 Belgium
6: 26830 Slovakia
7: 26830 Slovakia
8: 26830 Slovakia
9: 154694 Marc_Dutroux
10: 26830 Slovakia
11: 2596054 Topoľčany
12: 26830 Slovakia
13: 154694 Marc_Dutroux
14: 26830 Slovakia
15: 3343 Belgium
16: 154694 Marc_Dutroux
17: 26830 Slovakia
18: 5321 Czech_Republic
19: 154694 Marc_Dutroux
20: 154694 Marc_Dutroux
21: 2596054 Topoľčany
22: 2596054 Topoľčany
23: 154694 Marc_Dutroux
24: 26830 Slovakia
25: 154694 Marc_Dutroux
26: 3343 Belgium
27: 175228 Charleroi
28: 3343 Belgium
29: 154694 Marc_Dutroux
30: 26830 Slovakia
31: 26830 Slovakia
32: 26830 Slovakia
33: 3343 Belgium
34: 154694 Marc_Dutroux
35: 3343 Belgium
36: 3343 Belgium
37: 154694 Marc_Dutroux
1: 154694 Marc_Dutroux

[pem_pos=0, final_scores=0.57286537,
final_local_scores=0.6102697, pem=0.0, similarity=22.301378]
2: 26830 Slovakia
[pem_pos=1, final_scores=0.64089847,
final_local_scores=0.6245733, pem=-1.0527612,
similarity=28.453306]
3: 18933194 Bratislava
[pem_pos=0, final_scores=0.44124997,
final_local_scores=0.4303891, pem=0.0, similarity=16.933382]
4: 154694 Marc_Dutroux
[pem_pos=0, final_scores=0.59304965,
final_local_scores=0.63231796, pem=0.0, similarity=22.959343]
5: 3343 Belgium
[pem_pos=0, final_scores=0.6952318,
final_local_scores=0.6829531, pem=-0.12403572,
similarity=25.14492]
6: 26830 Slovakia
[pem_pos=1, final_scores=0.6290589,
final_local_scores=0.6116404, pem=-1.0527612,
similarity=28.06736]
7: 26830 Slovakia
[pem_pos=1, final_scores=0.6187687,
final_local_scores=0.6003999, pem=-1.0527612,
similarity=27.731924]
8: 26830 Slovakia
[pem_pos=1, final_scores=0.6431819,
final_local_scores=0.6270676, pem=-1.0527612,
similarity=28.527742]
9: 154694 Marc_Dutroux
[pem_pos=0, final_scores=0.66649073,
final_local_scores=0.7125412, pem=0.0, similarity=25.353365]
10: 26830 Slovakia
[pem_pos=1, final_scores=0.65136784,
final_local_scores=0.6360095, pem=-1.0527612,
similarity=28.794586]
11: 2596054 Topoľčany
[pem_pos=0, final_scores=0.4834525,
final_local_scores=0.52884245, pem=0.0, similarity=19.871426]
12: 26830 Slovakia
[pem_pos=0, final_scores=0.52395505,
final_local_scores=0.4968304, pem=-0.36602598,
similarity=20.90663]
13: 154694 Marc_Dutroux
[pem_pos=0, final_scores=0.63035375,
final_local_scores=0.67306703, pem=0.0, similarity=24.175375]
14: 26830 Slovakia
[pem_pos=1, final_scores=0.6194979,
final_local_scores=0.60119647, pem=-1.0527612,
similarity=27.755695]
15: 3343 Belgium
[pem_pos=1, final_scores=0.61464185,
final_local_scores=0.5949208, pem=-0.12403572,

```
similarity=22.517862]
16: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.6846012,
final_local_scores=0.7323241, pem=0.0, similarity=25.943726]
17: 26830 Slovakia
    [pem_pos=0, final_scores=0.5922036,
final_local_scores=0.57138157, pem=-0.36602598,
similarity=23.131386]
18: 5321 Czech_Republic
    [pem_pos=0, final_scores=0.42952043,
final_local_scores=0.39600307, pem=-0.27794555,
similarity=17.418745]
19: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.56713486,
final_local_scores=0.60401, pem=0.0, similarity=22.114576]
20: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.54829913,
final_local_scores=0.5834348, pem=0.0, similarity=21.500572]
21: 2596054 Topoľčany
    [pem_pos=0, final_scores=0.46722323,
final_local_scores=0.5111145, pem=0.0, similarity=19.342388]
22: 2596054 Topoľčany
    [pem_pos=0, final_scores=0.41885358,
final_local_scores=0.45827794, pem=0.0, similarity=17.76564]
23: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.65105546,
final_local_scores=0.6956805, pem=0.0, similarity=24.850206]
24: 26830 Slovakia
    [pem_pos=0, final_scores=0.55191267,
final_local_scores=0.52736986, pem=-0.36602598,
similarity=21.817987]
25: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.6499918,
final_local_scores=0.6945186, pem=0.0, similarity=24.815535]
26: 3343 Belgium
    [pem_pos=1, final_scores=0.6208064,
final_local_scores=0.60165465, pem=-0.12403572,
similarity=22.718813]
27: 175228 Charleroi
    [pem_pos=0, final_scores=0.39517838,
final_local_scores=0.41272384, pem=-0.2544922,
similarity=17.790184]
28: 3343 Belgium
    [pem_pos=1, final_scores=0.65377045,
final_local_scores=0.6376629, pem=-0.12403572,
similarity=23.793371]
29: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.6035597,
final_local_scores=0.6437986, pem=0.0, similarity=23.301947]
30: 26830 Slovakia
    [pem_pos=1, final_scores=0.6299194,
final_local_scores=0.61258036, pem=-1.0527612,
```

```
similarity=28.095413]
31: 26830 Slovakia
    [pem_pos=1, final_scores=0.6984967,
final_local_scores=0.68749064, pem=-1.0527612,
similarity=30.330885]
32: 26830 Slovakia
    [pem_pos=1, final_scores=0.66656905,
final_local_scores=0.65261453, pem=-1.0527612,
similarity=29.290112]
33: 3343 Belgium
    [pem_pos=0, final_scores=0.58736515,
final_local_scores=0.5651252, pem=-0.24180195,
similarity=22.269135]
34: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.63850373,
final_local_scores=0.68196964, pem=0.0, similarity=24.441048]
35: 3343 Belgium
    [pem_pos=0, final_scores=0.61348605,
final_local_scores=0.5936583, pem=-0.24180195,
similarity=23.120619]
36: 3343 Belgium
    [pem_pos=0, final_scores=0.6056001,
final_local_scores=0.58504415, pem=-0.24180195,
similarity=22.863554]
37: 154694 Marc_Dutroux
    [pem_pos=0, final_scores=0.59188324,
final_local_scores=0.63104385, pem=0.0, similarity=22.92132]
```

Sample: b'715_Knifeman'

Knifeman kills [1 1 Polish 1] beauty queen , wounds husband . [2 2 WARSAW 2] 1996-08-28 A man knifed to death international model [3 3 Agnieszka Kotlarska 3] outside her home in [4 4 Wroclaw 4] , western [5 5 Poland 5] , [6 6 Polish 6] television said on Wednesday . The man , who said he had once been engaged to her , first knifed [7 7 Kotlarska 7] 's husband in the leg , then stabbed her three times in the chest when she tried to intervene during the incident on Tuesday . She died in hospital . [8 8 Kotlarska 8] , who was 24 and had a three-year-old child , was Miss Poland in 1991 and went on to a [9 9 U.S.-based 9] modelling career that included working with [10 10 Italian 10] designer [11 11 Gian ni Versace 11] and [12 12 Vogue 12] magazine , the [13 13 Gaz eta Wyborcza 13] newspaper said . She had been due to fly on a [14 14 TWA 14] airliner which exploded near [15 15 New York 15] last month , but had cancelled her booking , the newspaper said . Her attacker , identified only as Jerzy L. , 36 , was arrested by police and will appear in court on Thursday morning , television reported . It said he had admitted the attack but had denied intending to kill [16 16 Ko tlarska 16] .

1: 22936 Poland
2: 32908 Warsaw

3: 4530919 Agnieszka_Kotlarska
4: 33603 Wrocław
5: 22936 Poland
6: 22936 Poland
7: 4530919 Agnieszka_Kotlarska
8: 4530919 Agnieszka_Kotlarska
9: 3434750 United_States
10: 14532 Italy
11: 318870 Gianni_Versace
12: 341903 Vogue_(magazine)
13: 657945 Gazeta_Wyborcza
14: 152449 Trans_World_Airlines
15: 645042 New_York_City
16: 4530919 Agnieszka_Kotlarska
1: 22936 Poland
[pem_pos=0, final_scores=0.71542305,
final_local_scores=0.7198644, pem=-0.81518245,
similarity=30.004993]
2: 32908 Warsaw
[pem_pos=0, final_scores=0.4888715,
final_local_scores=0.4920628, pem=-0.14300999,
similarity=19.55156]
3: 4530919 Agnieszka_Kotlarska
[pem_pos=0, final_scores=0.5188586,
final_local_scores=0.6090966, pem=0.0, similarity=22.26637]
4: 33603 Wrocław
[pem_pos=0, final_scores=0.48279983,
final_local_scores=0.49262166, pem=0.0, similarity=18.790525]
5: 22936 Poland
[pem_pos=0, final_scores=0.4959557,
final_local_scores=0.48012966, pem=-0.15333635,
similarity=19.251606]
6: 22936 Poland
[pem_pos=0, final_scores=0.55016345,
final_local_scores=0.5393434, pem=-0.81518245,
similarity=24.617884]
7: 4530919 Agnieszka_Kotlarska
[pem_pos=0, final_scores=0.5410295,
final_local_scores=0.6333149, pem=0.0, similarity=22.989094]
8: 4530919 Agnieszka_Kotlarska
[pem_pos=0, final_scores=0.6068608,
final_local_scores=0.70522565, pem=0.0, similarity=25.135054]
9: 3434750 United_States
[pem_pos=0, final_scores=0.4242925,
final_local_scores=0.44700032, pem=0.0, similarity=17.429094]
10: 14532 Italy
[pem_pos=0, final_scores=0.7288103,
final_local_scores=0.78438693, pem=-0.75948656,
similarity=31.62759]
11: 318870 Gianni_Versace
[pem_pos=0, final_scores=0.48676908,
final_local_scores=0.5902276, pem=0.0, similarity=21.703281]

```
12: 341903 Vogue_(magazine)
    [pem_pos=0, final_scores=0.3455249,
final_local_scores=0.39370018, pem=-0.16027622,
similarity=16.71012]
13: 657945 Gazeta_Wyborcza
    [pem_pos=0, final_scores=0.48147762,
final_local_scores=0.5619875, pem=0.0, similarity=20.860542]
14: 152449 Trans_World_Airlines
    [pem_pos=0, final_scores=0.3921427,
final_local_scores=0.48647934, pem=0.0, similarity=18.607227]
15: 645042 New_York_City
    [pem_pos=1, final_scores=0.44017613,
final_local_scores=0.44963616, pem=-1.7906687,
similarity=27.245687]
16: 4530919 Agnieszka_Kotlarska
    [pem_pos=0, final_scores=0.53565866,
final_local_scores=0.6274481, pem=0.0, similarity=22.814014]
```

Sample: b'716_Russian'

[1 1 Russian 1] shares slip in thin volume . [2 2 MOSCOW 2]
1996-08-28 Leading [3 3 Russian 3] shares edged down on
Wednesday in thin volume in the absence of [4 4 Western 4]
orders , traders said . The [5 5 Russian Trading System 5]
index of 21 issues fell 1.64 percent to 180.38 on volume of \$
4.38 million . " The market was extremely quiet today , some
profit-taking locally , no [6 6 Western 6] orders , " said
Nick Mokhoff , director of sales and trade at Alliance-
Menatep . " We are just a bit lower with a lot of inactivity
during the whole day . " Alexander Babayan , managing
director at CentrInvest Securities , said the volume of
orders was four to five times lower than a week ago . As
often , the most volume was in [7 7 UES 7] . " Unified is one
of the blue chips , which has more prospects than anybody
else does , because they have [8 8 ADRs 8] supposedly coming
up , " Mokhoff said . [9 9 UES 9] officials said last week
the board had not yet approved the final version of its
application to the U.S. Securities and Exchange Commission to
issue [10 10 ADRs 10] . [11 11 UES 11] fell to \$ 0.0817 from
\$ 0.0822 at Tuesday 's close with 8.90 million shares
changing hands . [12 12 Gazprom 12] was the loser of the day
with prices closing at \$ 0.300 , down from \$ 0.355 on Tuesday
and \$ 0.445 on Monday . Mokhoff said uncertainty about when [1
3 13 Gazprom 13] would issue [14 14 ADRs 14] and about
whether shares from the [15 15 Russian 15] market could be
converted into [16 16 ADRs 16] had hurt prices . " [17 17 Wes
tern 17] investors ... will be investing in [18 18 ADRs 18]
and I do not think people in [19 19 Russia 19] will be able
to come up with the money for the underlying shares to drive
the [20 20 Russian 20] shares to those levels , " Mokhoff
said . [21 21 Gazprom 21] has also tightened the rules
restricting shareholders ' rights to trade its shares . [22 22
Mosenergo 22] closed at at \$ 0.958 after \$ 0.966 , [23 23 Ros

telekom 23] fell to \$ 2.56 from \$ 2.58 and [24 24 LUKoil 24]
was \$ 9.82 after \$ 9.85 . -- Julie Tolkacheva , Moscow
Newsroom , +7095 941 8520

1: 25391 Russia
2: 19004 Moscow
3: 25391 Russia
4: 21208200 Western_world
5: 7191778 Russian_Trading_System
6: 21208200 Western_world
7: 1619047 RAO_UES
8: 412602 American_depository_receipt
9: 1619047 RAO_UES
10: 412602 American_depository_receipt
11: 1619047 RAO_UES
12: 320489 Gazprom
13: 320489 Gazprom
14: 412602 American_depository_receipt
15: 25391 Russia
16: 412602 American_depository_receipt
17: 21208200 Western_world
18: 412602 American_depository_receipt
19: 25391 Russia
20: 25391 Russia
21: 320489 Gazprom
22: 26006064 Mosenergo
23: 7208255 Rostelecom
24: 1027881 Lukoil

1: 25391 Russia
[pem_pos=1, final_scores=0.712488,
final_local_scores=0.7295145, pem=-1.0515022,
similarity=31.578112]

2: 19004 Moscow
[pem_pos=0, final_scores=0.62884384,
final_local_scores=0.6697758, pem=-0.106582455,
similarity=24.65677]

3: 25391 Russia
[pem_pos=1, final_scores=0.7450023,
final_local_scores=0.76503146, pem=-1.0515022,
similarity=32.63801]

4: 21208200 Western_world
[pem_pos=0, final_scores=0.3882844,
final_local_scores=0.43758327, pem=-1.5917408,
similarity=25.804205]

5: 7191778 Russian_Trading_System
[pem_pos=0, final_scores=0.51694256,
final_local_scores=0.59297204, pem=0.0, similarity=21.785181]

6: 21208200 Western_world
[pem_pos=0, final_scores=0.4719324,
final_local_scores=0.52895606, pem=-1.5917408,
similarity=28.53095]

7: 1619047 RAO_UES
[pem_pos=3, final_scores=0.5912508,

```
final_local_scores=0.63837355, pem=-1.0986123,
similarity=29.114477]
8: 412602 American_depository_receipt
    [pem_pos=0, final_scores=0.7742805,
final_local_scores=0.8584946, pem=0.0, similarity=29.708904]
9: 1619047 RA0_UES
    [pem_pos=3, final_scores=0.6040136,
final_local_scores=0.652315, pem=-1.0986123,
similarity=29.53052]
10: 412602 American_depository_receipt
    [pem_pos=0, final_scores=0.7948274,
final_local_scores=0.88093895, pem=0.0, similarity=30.378689]
11: 1619047 RA0_UES
    [pem_pos=3, final_scores=0.66350967,
final_local_scores=0.7173054, pem=-1.0986123,
similarity=31.469963]
12: 320489 Gazprom
    [pem_pos=0, final_scores=0.53330773,
final_local_scores=0.52664864, pem=0.0, similarity=19.80596]
13: 320489 Gazprom
    [pem_pos=0, final_scores=0.51519465,
final_local_scores=0.5068628, pem=0.0, similarity=19.215511]
14: 412602 American_depository_receipt
    [pem_pos=0, final_scores=0.7335789,
final_local_scores=0.8140344, pem=0.0, similarity=28.382124]
15: 25391 Russia
    [pem_pos=1, final_scores=0.81925136,
final_local_scores=0.84613717, pem=-1.0515022,
similarity=35.058365]
16: 412602 American_depository_receipt
    [pem_pos=0, final_scores=0.7589738,
final_local_scores=0.84177434, pem=0.0, similarity=29.209938]
17: 21208200 Western_world
    [pem_pos=0, final_scores=0.48876542,
final_local_scores=0.54734355, pem=-1.5917408,
similarity=29.07967]
18: 412602 American_depository_receipt
    [pem_pos=0, final_scores=1.0122762,
final_local_scores=1.1184688, pem=0.0, similarity=37.467052]
19: 25391 Russia
    [pem_pos=0, final_scores=0.7802656,
final_local_scores=0.8035513, pem=-0.20299491,
similarity=29.173204]
20: 25391 Russia
    [pem_pos=1, final_scores=0.71432775,
final_local_scores=0.7315241, pem=-1.0515022,
similarity=31.638083]
21: 320489 Gazprom
    [pem_pos=0, final_scores=0.3775633,
final_local_scores=0.35652146, pem=0.0, similarity=14.729024]
22: 26006064 Mosenergo
    [pem_pos=0, final_scores=0.45178974,
```

```
final_local_scores=0.55194396, pem=0.0, similarity=20.560822]
23: 7208255 Rostelecom
    [pem_pos=0, final_scores=0.5982412,
final_local_scores=0.64270717, pem=0.0, similarity=23.269377]
24: 1027881 Lukoil
    [pem_pos=0, final_scores=0.6323833,
final_local_scores=0.6771887, pem=0.0, similarity=24.298374]
```

Sample: b'717_Albania'

[1 1 Albania 1] asks [2 2 Greece 2] to explain deportations . [3 3 TIRANA 3] 1996-08-28 [4 4 Albania 4] asked [5 5 Greece 5] on Wednesday to explain why it was deporting more [6 6 Albanian 6] immigrants , Foreign Minister Tritan Shehu said . The [7 7 Albanian 7] daily [8 8 Koha Jone 8] reported earlier that [9 9 Greece 9] had deported about 5,000 [10 10 Albanians 10] in the last five days . " The [11 1 1 Foreign Ministry 11] is trying to find out from the [12 12 G reek 12] embassy why [13 13 Albanian 13] refugees have been deported from [14 14 Greece 14] , " Shehu told [1 Reuters 1] . [15 15 Athens 15] and [16 16 Tirana 16] signed an accord in May to legalise the status of [17 17 Albanian 17] immigrant workers , estimated at 350,000 , and remove a long-standing stumbling block in relations between the two [18 18 Balkan 18] neighbours .

```
1: 738 Albania
2: 12108 Greece
3: 31075 Tirana
4: 738 Albania
5: 12108 Greece
6: 738 Albania
7: 738 Albania
8: 27830610 Koha_Jonë
9: 12108 Greece
10: 2927 Albanians
11: 317153 Foreign_minister
12: 12108 Greece
13: 738 Albania
14: 12108 Greece
15: 1216 Athens
16: 31075 Tirana
17: 738 Albania
18: 4829 Balkans
1: 738 Albania
    [pem_pos=0, final_scores=0.5859278,
final_local_scores=0.5312746, pem=-0.41162592,
similarity=22.182493]
2: 12108 Greece
    [pem_pos=0, final_scores=0.99242216,
final_local_scores=0.98827046, pem=-0.27607638,
similarity=35.083023]
3: 31075 Tirana
    [pem_pos=0, final_scores=0.4137352,
```

```
final_local_scores=0.38743055, pem=0.0, similarity=15.651413]
4: 738 Albania
    [pem_pos=0, final_scores=0.60684973,
final_local_scores=0.55412865, pem=-0.41162592,
similarity=22.864502]
5: 12108 Greece
    [pem_pos=0, final_scores=0.7816219,
final_local_scores=0.7580032, pem=-0.27607638,
similarity=28.211386]
6: 738 Albania
    [pem_pos=2, final_scores=0.6614464,
final_local_scores=0.6137673, pem=-1.5560378,
similarity=30.867731]
7: 738 Albania
    [pem_pos=2, final_scores=0.6081212,
final_local_scores=0.55551755, pem=-1.5560378,
similarity=29.12944]
8: 27830610 Koha_Jonë
    [pem_pos=0, final_scores=0.5675069,
final_local_scores=0.70347446, pem=0.0, similarity=25.082794]
9: 12108 Greece
    [pem_pos=0, final_scores=0.8947125,
final_local_scores=0.8815376, pem=-0.27607638,
similarity=31.8979]
10: 2927 Albanians
    [pem_pos=0, final_scores=0.31425452,
final_local_scores=0.2817062, pem=-0.028324477,
similarity=12.650421]
11: 317153 Foreign_minister
    [pem_pos=0, final_scores=0.26435316,
final_local_scores=0.30367625, pem=-0.35490593,
similarity=15.082051]
12: 12108 Greece
    [pem_pos=1, final_scores=0.9023934,
final_local_scores=0.8899277, pem=-2.3247876,
similarity=43.289486]
13: 738 Albania
    [pem_pos=2, final_scores=0.6635439,
final_local_scores=0.6160584, pem=-1.5560378,
similarity=30.936102]
14: 12108 Greece
    [pem_pos=0, final_scores=0.88914186,
final_local_scores=0.87545234, pem=-0.27607638,
similarity=31.716305]
15: 1216 Athens
    [pem_pos=0, final_scores=0.4146291,
final_local_scores=0.4117782, pem=-0.28831622,
similarity=17.945906]
16: 31075 Tirana
    [pem_pos=0, final_scores=0.4025877,
final_local_scores=0.37525362, pem=0.0, similarity=15.28803]
17: 738 Albania
```

```
[pem_pos=2, final_scores=0.71077436,
final_local_scores=0.6676505, pem=-1.5560378,
similarity=32.47571]
18: 4829 Balkans
    [pem_pos=0, final_scores=0.4619301,
final_local_scores=0.45590067, pem=-0.13899443,
similarity=18.450571]
1: 18998750 Reuters
    [pem_pos=0, final_scores=0.21233754,
final_local_scores=0.26735628, pem=0.0, similarity=12.068157]
```

Sample: b'718_Bulgarians'

```
[1 1 Bulgarians 1] recover 75 pct of confiscated land . [2 2
SOFIA 2] 1996-08-28 [3 3 Bulgaria 3] has restored ownership
rights to pre-communist private owners of 75 percent of the
arable land or around four million hectares , an Agriculture
Ministry official said on Wednesday . " So far 75 percent has
been returned with the land restitution almost completed in
some regions like Southeastern Bulgaria but lagging behind in
other areas , predominantly in the mountains , " the official
told a news conference . The ministry has said that it
planned to return 96.6 percent of the arable land or 5.2
million hectares to its original owners by the end of this
year . A land reform act passed four years ago abolished
Soviet-style collective farms , allowing the return of 5.4
million hectares to original owners or their heirs . -- Sofia
Newsroom , ( +359-2 ) 981 8569
```

```
1: 151876 Bulgarians
2: 57644 Sofia
3: 3415 Bulgaria
1: 151876 Bulgarians
    [pem_pos=0, final_scores=0.2780916,
final_local_scores=0.24981466, pem=-0.15953687,
similarity=12.412267]
2: 57644 Sofia
    [pem_pos=0, final_scores=0.5189908,
final_local_scores=0.5000192, pem=-0.10680699,
similarity=19.592115]
3: 3415 Bulgaria
    [pem_pos=0, final_scores=0.7110754,
final_local_scores=0.6964748, pem=-0.31727514,
similarity=26.599302]
```

Sample: b'719_German'

```
[1 1 German 1] chancellor to meet [2 2 Yeltsin 2] Sept 7 - [3
3 Interfax 3] . [4 4 MOSCOW 4] 1996-08-28 [5 5 German 5]
Chancellor [6 6 Helmut Kohl 6] , who spoke to President [7 7 B
oris Yeltsin 7] by telephone on Wednesday , plans a trip to [8
8 Moscow 8] on September 7 and will visit [9 9 Yeltsin 9] at
his vacation home near [10 10 Moscow 10] , [11 11 Interfax 11]
news agency said . [12 12 Interfax 12] , quoting [13 13 Yelts
in 13] press secretary [14 14 Sergei Yastrzhembsky 14] , said
```

[15 15 Yeltsin 15] and [16 16 Kohl 16] had discussed bilateral relations and international issues on the telephone . [17 17 Yeltsin 17] had told [18 18 Kohl 18] about efforts to find a political solution to the conflict in [19 19 Russia 19] 's breakaway [20 20 Chechnya 20] region , [21 21 Interfax 21] added . [22 22 Yeltsin 22] , who left on vacation on Monday , is staying at an exclusive private hunting lodge some 100 km (60 miles) from [23 23 Moscow 23] . [24 24 Germany 24] has been one of the loudest critics of [25 25 Russia 25] 's military intervention in [26 26 Chechnya 26] , a 20-month-old conflict in which tens of thousands of people have been killed .

1: 11867 Germany
2: 23278520 Boris_Yeltsin
3: 4921404 Interfax
4: 19004 Moscow
5: 11867 Germany
6: 41896 Helmut_Kohl
7: 23278520 Boris_Yeltsin
8: 19004 Moscow
9: 23278520 Boris_Yeltsin
10: 19004 Moscow
11: 4921404 Interfax
12: 4921404 Interfax
13: 23278520 Boris_Yeltsin
14: 13411146 Sergey_Yastrzhembsky
15: 23278520 Boris_Yeltsin
16: 41896 Helmut_Kohl
17: 23278520 Boris_Yeltsin
18: 41896 Helmut_Kohl
19: 25391 Russia
20: 6095 Chechnya
21: 4921404 Interfax
22: 23278520 Boris_Yeltsin
23: 19004 Moscow
24: 11867 Germany
25: 25391 Russia
26: 6095 Chechnya
1: 11867 Germany
[pem_pos=2, final_scores=0.77048045,
final_local_scores=0.8244608, pem=-1.0465823,
similarity=34.384747]
2: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.5591505,
final_local_scores=0.5237953, pem=0.0, similarity=19.72081]
3: 4921404 Interfax
[pem_pos=0, final_scores=0.42946678,
final_local_scores=0.44401747, pem=0.0, similarity=17.34008]
4: 19004 Moscow
[pem_pos=0, final_scores=0.7021046,
final_local_scores=0.70065945, pem=-0.106582455,
similarity=25.578402]

5: 11867 Germany
[pem_pos=1, final_scores=0.7490342,
final_local_scores=0.80103403, pem=-1.0465823,
similarity=33.685646]
6: 41896 Helmut_Kohl
[pem_pos=0, final_scores=0.865267,
final_local_scores=0.9078173, pem=0.0, similarity=31.180794]
7: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.70101345,
final_local_scores=0.678759, pem=0.0, similarity=24.345234]
8: 19004 Moscow
[pem_pos=0, final_scores=0.64268285,
final_local_scores=0.6357502, pem=-0.106582455,
similarity=23.641378]
9: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.51743585,
final_local_scores=0.4782284, pem=0.0, similarity=18.361002]
10: 19004 Moscow
[pem_pos=0, final_scores=0.5378753,
final_local_scores=0.5212639, pem=-0.106582455,
similarity=20.22488]
11: 4921404 Interfax
[pem_pos=0, final_scores=0.7337251,
final_local_scores=0.77637345, pem=0.0, similarity=27.258245]
12: 4921404 Interfax
[pem_pos=0, final_scores=0.55761635,
final_local_scores=0.5840014, pem=0.0, similarity=21.51748]
13: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.5648982,
final_local_scores=0.53007376, pem=0.0, similarity=19.90817]
14: 13411146 Sergey_Yastrzhembsky
[pem_pos=0, final_scores=0.6375107,
final_local_scores=0.71836656, pem=0.0, similarity=25.527205]
15: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.63715595,
final_local_scores=0.6090044, pem=0.0, similarity=22.263618]
16: 41896 Helmut_Kohl
[pem_pos=0, final_scores=0.907353,
final_local_scores=0.9614927, pem=0.0, similarity=32.782574]
17: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.6991692,
final_local_scores=0.67674446, pem=0.0, similarity=24.285118]
18: 41896 Helmut_Kohl
[pem_pos=0, final_scores=0.83705693,
final_local_scores=0.884705, pem=0.0, similarity=30.491074]
19: 25391 Russia
[pem_pos=0, final_scores=0.89172393,
final_local_scores=0.9373978, pem=-0.20299491,
similarity=33.16745]
20: 6095 Chechnya
[pem_pos=0, final_scores=0.5316617,
final_local_scores=0.5436823, pem=0.0, similarity=20.314276]

```
21: 4921404 Interfax
    [pem_pos=0, final_scores=0.56228065,
final_local_scores=0.5890964, pem=0.0, similarity=21.669525]
22: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.69534796,
final_local_scores=0.6725703, pem=0.0, similarity=24.160551]
23: 19004 Moscow
    [pem_pos=0, final_scores=0.5422419,
final_local_scores=0.5260337, pem=-0.106582455,
similarity=20.36722]
24: 11867 Germany
    [pem_pos=0, final_scores=0.56996185,
final_local_scores=0.60542476, pem=-0.15529114,
similarity=23.001293]
25: 25391 Russia
    [pem_pos=0, final_scores=0.9120138,
final_local_scores=0.9595613, pem=-0.20299491,
similarity=33.828854]
26: 6095 Chechnya
    [pem_pos=0, final_scores=0.49509746,
final_local_scores=0.50374144, pem=0.0, similarity=19.122362]
```

Sample: b'720_Exiled'

Exiled Bosnians protest confusing voting rules . Duncan Shiels [1 1 NAGYATAD 1] , [2 2 Hungary 2] 1996-08-28 Bosnian refugees in [3 3 Hungary 3] , the first to vote last weekend in their country 's first post-war election , found the rules confusing and some had no idea who they voted for , refugees and officials said on Wednesday . " For the most part they really did n't understand what was going on , " the director of the [4 4 Nagyatad 4] camp Lajos Horvath told [5 5 Reuters 5] on Wednesday . " It was confusing , they had no experience of voting , many of the refugees are only semi-literate and none of them knew anything about the candidates , " he said . " They just voted along ethnic lines where they could . " The Bosnian election is set for September 14 and voting began on Wednesday for most of the 600,000 Bosnians living abroad . The Organisation for Security and Cooperation in Europe , which is running the election , allowed the ballot to be held on Sunday in four [6 6 Hungarian 6] refugee camps . Some [7 7 Moslem 7] refugees among the 385 registered voters [8 8 Hungry 8] 's largest camp [9 9 Nagyatad 9] have written to the OSCE complaining that they were unable to vote in contests for the president or assembly of the Bosnian-Croat Federation , where most [10 10 Moslems 10] live . This was because their pre-war homes are now in the Serb-controlled territory , so they were voting for the national assembly of the Republika Srbska , most of the candidates for which are [1 1 11 Serbs 11] . For the foreign powers which back last year 's [12 12 Dayton 12] peace agreement , the main point of the election rules is that by voting as though they were still in their pre-war homes , Bosnians should override the effects of

ethnic cleansing and reassert the concept of a single multi-ethnic state . But Adem Hodzic , one of the refugees who signed the letter of complaint , told [13 13 Reuters 13] : " We only realised after voting that we were being denied the rights of other Bosnian Moslems to choose our president . This vote seals the division of my country . " Under the election rules citizens vote for a three-man presidency and House of Representatives for all Bosnia-Hercegovina and for assemblies and cantonal seats in either the Moslem-Croat Federation or the Republika Srbska . On Sunday voters in [14 1 4 Hungary 14] also cast ballots for municipal councils but these will be invalidated following the cancellation of local elections by the OSCE on Tuesday . Husein Micijevic , who also signed the letter , alleged that elderly voters were directed who to vote for by [15 15 Hungarian 15] translators who stood in the polling booth to help them . " Probably 100 refugees were shown where to put their cross , " he said . Seventy eight-year-old Mandolina Zelic , a Bosnian Croat who has spent the last five years in [16 16 Nagyatad 16] , told [1 7 17 Reuters 17] she had cast her ballot because she 'd been told to by the camp authorities but had no idea who she voted for . " At first the organisers would n't let anyone help me but when they saw I did n't understand a young translator ringed the names I had to mark , " she said . " I do n't know who I voted for . " Maria Szabo of the [18 18 Hungarian 18] office organising the elections on behalf of the OSCE told [19 19 Reuters 19] on Wednesday her office was studying the letter but said they had followed OSCE instructions very carefully . " The envelopes , each with the five different voting slips , were sealed until voting and had written instructions on how to vote , " she said . " But of course those who could not read had to be shown . "

- 1: 11274042 Nagyatd
- 2: 13275 Hungary
- 3: 13275 Hungary
- 4: 11274042 Nagyatd
- 5: 18998750 Reuters
- 6: 13275 Hungary
- 7: 6037917 Islam
- 8: 13275 Hungary
- 9: 11274042 Nagyatd
- 10: 6037917 Islam
- 11: 29265 Serbia
- 12: 8253 Dayton,_Ohio
- 13: 18998750 Reuters
- 14: 13275 Hungary
- 15: 13275 Hungary
- 16: 11274042 Nagyatd
- 17: 18998750 Reuters
- 18: 13275 Hungary
- 19: 18998750 Reuters
- 1: 11274042 Nagyatd

```
[pem_pos=0, final_scores=0.4216786,
final_local_scores=0.5161748, pem=0.0, similarity=19.493399]
2: 13275 Hungary
[pem_pos=0, final_scores=0.6574363,
final_local_scores=0.63398623, pem=-0.32577527,
similarity=24.780745]
3: 13275 Hungary
[pem_pos=0, final_scores=0.68257743,
final_local_scores=0.66144913, pem=-0.32577527,
similarity=25.600292]
4: 11274042 Nagyatd
[pem_pos=0, final_scores=0.43017817,
final_local_scores=0.5254593, pem=0.0, similarity=19.770466]
5: 18998750 Reuters
[pem_pos=0, final_scores=0.43402767,
final_local_scores=0.4756844, pem=0.0, similarity=18.285084]
6: 13275 Hungary
[pem_pos=1, final_scores=0.6981319,
final_local_scores=0.67844003, pem=-1.0319586,
similarity=29.94767]
7: 6037917 Islam
[pem_pos=1, final_scores=0.4336608,
final_local_scores=0.51033664, pem=-1.2150909,
similarity=25.927029]
8: 13275 Hungary
[pem_pos=0, final_scores=0.6246607,
final_local_scores=0.5981839, pem=-0.32577527,
similarity=23.71233]
9: 11274042 Nagyatd
[pem_pos=0, final_scores=0.52017015,
final_local_scores=0.6237619, pem=0.0, similarity=22.704012]
10: 6037917 Islam
[pem_pos=0, final_scores=0.5763679,
final_local_scores=0.66622245, pem=0.0, similarity=23.971119]
11: 29265 Serbia
[pem_pos=4, final_scores=0.34649187,
final_local_scores=0.3311138, pem=-2.3277044,
similarity=26.629227]
12: 8253 Dayton,_Ohio
[pem_pos=0, final_scores=0.2597196,
final_local_scores=0.3149408, pem=-0.9297534,
similarity=18.54432]
13: 18998750 Reuters
[pem_pos=0, final_scores=0.3631118,
final_local_scores=0.3982196, pem=0.0, similarity=15.97338]
14: 13275 Hungary
[pem_pos=0, final_scores=0.57291275,
final_local_scores=0.54165715, pem=-0.32577527,
similarity=22.02546]
15: 13275 Hungary
[pem_pos=1, final_scores=0.4970911,
final_local_scores=0.45883352, pem=-1.0319586,
```

```
similarity=23.394173]
16: 11274042 Nagyatd
    [pem_pos=0, final_scores=0.48275393,
final_local_scores=0.58289033, pem=0.0, similarity=21.484324]
17: 18998750 Reuters
    [pem_pos=0, final_scores=0.40518552,
final_local_scores=0.4441787, pem=0.0, similarity=17.34489]
18: 13275 Hungary
    [pem_pos=1, final_scores=0.6575228,
final_local_scores=0.6340807, pem=-1.0319586,
similarity=28.623898]
19: 18998750 Reuters
    [pem_pos=0, final_scores=0.40388972,
final_local_scores=0.44276327, pem=0.0, similarity=17.302652]
```

Sample: b'721_Estonia'

[1 1 Estonia 1] presidential race next round on Sept 20 . [2 2 TALLINN 2] 1996-08-28 [3 3 Estonia 3] will hold the next round of an inconclusive state presidential race on September 20 , parliamentary officers of the Baltic state ruled on Wednesday . This comes after three votes in the 101-strong parliament on Monday and Tuesday failed to give either incumbent [4 4 Lennart Meri 4] or rival candidate [5 5 Arnold Ruutel 5] the necessary 68 votes for a clear mandate . The outcome was a rebuff for [6 6 Meri 6] , failing three times to win backing in his bid for a second term as head of state of the former [7 7 Soviet 7] republic . Parliament 's press officer told [8 8 Reuters 8] that Speaker [9 9 Toomas Savi 9] will convene an electoral college involving 101 MPs and and 273 local government representatives on September 20 . Both [1 0 10 Meri 10] , 67 , and [11 11 Ruutel 11] , 68 will automatically be listed as candidates but the election will also be open to new nominations with the backing of any 21 members of the college . The winner has to secure a majority from the college within two rounds of voting otherwise the election will go back before the parliament .

```
1: 28222445 Estonia
2: 31577 Tallinn
3: 28222445 Estonia
4: 379531 Lennart_Meri
5: 341466 Arnold_Rtel
6: 379531 Lennart_Meri
7: 26779 Soviet_Union
8: 18998750 Reuters
9: 1473586 Toomas_Savi
10: 379531 Lennart_Meri
11: 341466 Arnold_Rtel
1: 28222445 Estonia
    [pem_pos=0, final_scores=0.4826122,
final_local_scores=0.49482107, pem=-0.4106629,
similarity=21.089409]
2: 31577 Tallinn
```

```
[pem_pos=0, final_scores=0.4833185,
final_local_scores=0.49287736, pem=0.0, similarity=18.798157]
3: 28222445 Estonia
    [pem_pos=0, final_scores=0.5087993,
final_local_scores=0.52342653, pem=-0.4106629,
similarity=21.943054]
4: 379531 Lennart_Meri
    [pem_pos=0, final_scores=1.0706477,
final_local_scores=1.0788884, pem=0.0, similarity=36.285896]
5: 341466 Arnold_Rüütel
    [pem_pos=0, final_scores=0.6821585,
final_local_scores=0.71882415, pem=0.0, similarity=25.54086]
6: 379531 Lennart_Meri
    [pem_pos=0, final_scores=0.9268603,
final_local_scores=0.93042225, pem=0.0, similarity=31.85537]
7: 26779 Soviet_Union
    [pem_pos=0, final_scores=0.41704923,
final_local_scores=0.42529285, pem=-0.097171314,
similarity=17.309732]
8: 18998750 Reuters
    [pem_pos=0, final_scores=0.36732376,
final_local_scores=0.45977592, pem=0.0, similarity=17.810343]
9: 1473586 Toomas_Savi
    [pem_pos=0, final_scores=0.35237432,
final_local_scores=0.46108449, pem=0.0, similarity=17.849394]
10: 379531 Lennart_Meri
    [pem_pos=0, final_scores=0.9054131,
final_local_scores=0.9069944, pem=0.0, similarity=31.156235]
11: 341466 Arnold_Rüütel
    [pem_pos=0, final_scores=0.53236765,
final_local_scores=0.55520046, pem=0.0, similarity=20.658003]
```

Sample: b'722_Lebed'

[1 1 Lebed 1] likely to fail on [2 2 Chechnya 2] - [3 3 Polish 3] minister . [4 4 WARSAW 4] 1996-08-28 [5 5 Russian 5] security chief Aleksander Lebed faces an almost impossible task in [6 6 Chechnya 6] and is likely to be sidelined , [7 7 Polish 7] Foreign Minister [8 8 Dariusz Rosati 8] was reported as saying on Wednesday . According to best-selling daily [9 9 Gazeta Wyborcza 9] , [10 10 Rosati 10] told the [11 11 Polish 11] parliament 's foreign affairs committee on Tuesday that the fact [12 12 Lebed 12] had been charged with resolving the conflict in [13 13 Chechnya 13] showed he would be marginalised . " It is almost impossible to gain success in this , " it quoted [14 14 Rosati 14] as saying during a committee debate . " [15 15 Lebed 15] has no diplomatic experience . [16 16 Yeltsin 16] sent him there to compromise him . This tactical manoeuvre also shows that in the ruling circle there is no unity of action , " he said . [17 17 Lebed 17] , who has arranged a military truce with separatist rebels in the southern [18 18 Russia 18] region , was in [19 19 Moscow 19] this week seeking support for a deal on [20 20 Ch

echnya 20] 's political status . But [21 21 Russian 21] President [22 22 Boris Yeltsin 22] has seemed unwilling to meet his envoy and went on holiday on Monday . [23 23 Gazeta Wyborcza 23] quoted [24 24 Rosati 24] as saying [25 25 Yeltsin 25] was very ill and effectively on leave , but for now retained control in [26 26 Russia 26] although matters were passing into the hands of his close collaborators . [27 27 Rosati 27] said [28 28 Russia 28] 's July polls , in which [29 29 Yeltsin 29] won re-election , showed democracy had passed an important test and the [30 30 Russian 30] people had chosen the path of further reforms . But he said a power struggle in [31 31 Russia 31] 's ruling circles could not be ruled out , which could harm further reforms . He expressed concern over problems in the [32 32 Russian 32] economy , saying this could lead to social unrest , the daily reported . On 1 Moscow 's foreign policy , [33 33 Rosati 33] said it had changed its stance on [34 34 NATO 34] 's eastward expansion and was preparing itself for [35 35 Poland 35] 's inevitable entry into the 2 Western alliance . He also reportedly criticised [36 36 Russian 36] Foreign Minister [37 37 Yevgeny Primakov 37] , saying his style of work resembled that of the Soviet-era 1970s and 1980s .

1: 431579 Alexander_Lebed
2: 6095 Chechnya
3: 22936 Poland
4: 32908 Warsaw
5: 25391 Russia
6: 6095 Chechnya
7: 22936 Poland
8: 776028 Dariusz_Rosati
9: 657945 Gazeta_Wyborcza
10: 776028 Dariusz_Rosati
11: 22936 Poland
12: 431579 Alexander_Lebed
13: 6095 Chechnya
14: 776028 Dariusz_Rosati
15: 431579 Alexander_Lebed
16: 23278520 Boris_Yeltsin
17: 431579 Alexander_Lebed
18: 25391 Russia
19: 19004 Moscow
20: 6095 Chechnya
21: 25391 Russia
22: 23278520 Boris_Yeltsin
23: 657945 Gazeta_Wyborcza
24: 776028 Dariusz_Rosati
25: 23278520 Boris_Yeltsin
26: 25391 Russia
27: 776028 Dariusz_Rosati
28: 25391 Russia
29: 23278520 Boris_Yeltsin
30: 25391 Russia

31: 25391 Russia
32: 25391 Russia
33: 776028 Dariusz_Rosati
34: 21133 NATO
35: 22936 Poland
36: 25391 Russia
37: 520754 Yevgeny_Primakov

1: 25391 Russia
[winner: 19004 Moscow, pem_pos=0,
final_scores=-0.27961135, final_local_scores=-0.36229086,
pem=-0.106582455, similarity=-6.1421576]
RECALL MISS: 25391 Russia]

2: 21133 NATO
[winner: 21208200 Western_world, pem_pos=0,
final_scores=-0.056402996, final_local_scores=-0.056134373,
pem=-1.5917408, similarity=11.070685]
RECALL MISS: 21133 NATO]

1: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.5521637,
final_local_scores=0.5952773, pem=-0.2123326,
similarity=23.008673]

2: 6095 Chechnya
[pem_pos=0, final_scores=0.5396519,
final_local_scores=0.5610287, pem=0.0, similarity=20.831928]

3: 22936 Poland
[pem_pos=1, final_scores=0.6189967,
final_local_scores=0.6197266, pem=-0.81518245,
similarity=27.016678]

4: 32908 Warsaw
[pem_pos=0, final_scores=0.48006338,
final_local_scores=0.49086148, pem=-0.14300999,
similarity=19.515709]

5: 25391 Russia
[pem_pos=2, final_scores=0.9541121,
final_local_scores=0.99063474, pem=-1.0515022,
similarity=39.370464]

6: 6095 Chechnya
[pem_pos=0, final_scores=0.47546226,
final_local_scores=0.4909113, pem=0.0, similarity=18.739485]

7: 22936 Poland
[pem_pos=0, final_scores=0.6054707,
final_local_scores=0.60495156, pem=-0.81518245,
similarity=26.575764]

8: 776028 Dariusz_Rosati
[pem_pos=0, final_scores=0.75574,
final_local_scores=0.8316314, pem=0.0, similarity=28.907253]

9: 657945 Gazeta_Wyborcza
[pem_pos=0, final_scores=0.43265754,
final_local_scores=0.49204963, pem=0.0, similarity=18.773455]

10: 776028 Dariusz_Rosati
[pem_pos=0, final_scores=1.1994828,
final_local_scores=1.3237777, pem=0.0, similarity=43.59388]

11: 22936 Poland
[pem_pos=1, final_scores=0.63990194,
final_local_scores=0.64256245, pem=-0.81518245,
similarity=27.698149]
12: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.6720448,
final_local_scores=0.7262292, pem=-0.2123326,
similarity=26.916538]
13: 6095 Chechnya
[pem_pos=0, final_scores=0.52305096,
final_local_scores=0.5428947, pem=0.0, similarity=20.290773]
14: 776028 Dariusz_Rosati
[pem_pos=0, final_scores=1.1021436,
final_local_scores=1.2174494, pem=0.0, similarity=40.42083]
15: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.5328168,
final_local_scores=0.57414377, pem=-0.2123326,
similarity=22.378006]
16: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.6454182,
final_local_scores=0.6538279, pem=0.0, similarity=23.601242]
17: 431579 Alexander_Lebed
[pem_pos=0, final_scores=0.6399224,
final_local_scores=0.6911403, pem=-0.2123326,
similarity=25.869415]
18: 25391 Russia
[pem_pos=0, final_scores=0.78096443,
final_local_scores=0.8014971, pem=-0.20299491,
similarity=29.111902]
19: 19004 Moscow
[pem_pos=0, final_scores=0.67298645,
final_local_scores=0.682529, pem=-0.106582455,
similarity=25.037352]
20: 6095 Chechnya
[pem_pos=0, final_scores=0.5324353,
final_local_scores=0.5531457, pem=0.0, similarity=20.596684]
21: 25391 Russia
[pem_pos=1, final_scores=0.8833266,
final_local_scores=0.9133124, pem=-1.0515022,
similarity=37.06301]
22: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.7508386,
final_local_scores=0.76898354, pem=0.0, similarity=27.037716]
23: 657945 Gazeta_Wyborcza
[pem_pos=0, final_scores=0.3747297,
final_local_scores=0.42877227, pem=0.0, similarity=16.885132]
24: 776028 Dariusz_Rosati
[pem_pos=0, final_scores=1.2577857,
final_local_scores=1.3874648, pem=0.0, similarity=45.494427]
25: 23278520 Boris_Yeltsin
[pem_pos=0, final_scores=0.5881937,
final_local_scores=0.5913188, pem=0.0, similarity=21.735846]

```
26: 25391 Russia
    [pem_pos=0, final_scores=0.7790275,
final_local_scores=0.79938143, pem=-0.20299491,
similarity=29.048767]
27: 776028 Dariusz_Rosati
    [pem_pos=0, final_scores=1.1669543,
final_local_scores=1.2882452, pem=0.0, similarity=42.533516]
28: 25391 Russia
    [pem_pos=0, final_scores=0.928806,
final_local_scores=0.96299165, pem=-0.20299491,
similarity=33.93122]
29: 23278520 Boris_Yeltsin
    [pem_pos=0, final_scores=0.52633333,
final_local_scores=0.5237457, pem=0.0, similarity=19.71933]
30: 25391 Russia
    [pem_pos=2, final_scores=0.8065856,
final_local_scores=0.82948446, pem=-1.0515022,
similarity=34.561417]
31: 25391 Russia
    [pem_pos=0, final_scores=0.8709057,
final_local_scores=0.89974433, pem=-0.20299491,
similarity=32.043797]
32: 25391 Russia
    [pem_pos=2, final_scores=0.81081754,
final_local_scores=0.83410716, pem=-1.0515022,
similarity=34.699368]
33: 776028 Dariusz_Rosati
    [pem_pos=0, final_scores=1.2241001,
final_local_scores=1.3506684, pem=0.0, similarity=44.39635]
34: 21133 NATO
    [pem_pos=0, final_scores=0.41817796,
final_local_scores=0.43518567, pem=0.0, similarity=17.07652]
35: 22936 Poland
    [pem_pos=0, final_scores=0.5575592,
final_local_scores=0.55261546, pem=-0.15333635,
similarity=21.414726]
36: 25391 Russia
    [pem_pos=2, final_scores=0.947267,
final_local_scores=0.98315746, pem=-1.0515022,
similarity=39.147327]
37: 520754 Yevgeny_Primakov
    [pem_pos=0, final_scores=0.7401032,
final_local_scores=0.8090393, pem=0.0, similarity=28.233059]
```

Sample: b'723_Romania'

```
[1 1 Romania 1] state budget soars in June . [2 2 BUCHAREST 2
] 1996-08-28 [3 3 Romania 3] 's state budget deficit jumped
sharply in June to 1,242.9 billion lei for the January-June
period from 596.5 billion lei in January-May , official data
showed on Wednesday . Six-month expenditures stood at 9.50
trillion lei , up from 7.56 trillion lei at end-May , with
education and health spending accounting for 31.6 percent of
```

state expenses and economic subsidies and support taking some 26 percent . January-June revenues went up to 8.26 trillion lei from 6.96 trillion lei in the first five months this year . [4 4 Romania 4] 's government is expected to revise the 1996 budget on Wednesday to bring it into line with higher inflation , new wage and pension indexations and costs of energy imports that have pushed up the state deficit . Under the revised version state spending is expected to rise by some 566 billion lei . No new deficit forecast has been issued so far . In July the government gave a 6.0-percent wage and pension indexation to cover energy , fuel and bread price increases , which quickened inflation to 7.5 percent last month . In the original state budget , approved in March , revenues were envisaged at around 16.98 trillion lei and expenditures 20.17 trillion lei for 1996 . The state budget deficit was originally forecast to be 3.19 trillion lei for the whole year . On Wednesday , the leu 's official rate was 3,161 to the dollar . -- Bucharest Newsroom 40-1 3120264

1: 25445 Romania
2: 36877 Bucharest
3: 25445 Romania
4: 25445 Romania
1: 25445 Romania
[pem_pos=0, final_scores=0.7860063,
final_local_scores=0.736822, pem=-0.21241479,
similarity=27.233095]
2: 36877 Bucharest
[pem_pos=0, final_scores=0.4885764,
final_local_scores=0.4625228, pem=-0.01781679,
similarity=17.989206]
3: 25445 Romania
[pem_pos=0, final_scores=0.81407374,
final_local_scores=0.7674814, pem=-0.21241479,
similarity=28.148033]
4: 25445 Romania
[pem_pos=0, final_scores=0.771738,
final_local_scores=0.721236, pem=-0.21241479,
similarity=26.767979]

Sample: b'724_Costa'
[1 1 Costa Rica 1] says [2 2 Dutch 2] pair kidnapped by [3 3
Nicaraguans 3] . [4 4 SAN JOSE 4] , [5 5 Costa Rica 5]
1996-08-28 The [6 6 Costa Rican 6] government said on
Wednesday that a [7 7 Dutch 7] couple abducted over the
weekend from a tree farm in northern [8 8 Costa Rica 8] was
kidnapped by former [9 9 Nicaraguan 9] guerrillas . " Even
though it 's an act of common delinquency , the case could
take a difficult turn because former [10 10 Nicaraguan 10]
guerrillas are involved , " Security Minister Bernardo Arce
told reporters . Earlier this year , a [11 11 German 11]
tourist and a [12 12 Swiss 12] tour guide were kidnapped from

the same general area in northern [13 13 Costa Rica 13] near the [14 14 Nicaraguan 14] border . They were held for 71 days before relatives paid a ransom to free them . Two [15 15 Nica raguan 15] former guerrillas have been arrested in the case . Because of the apparent threat to foreigners in [16 16 Costa R ica 16] near the [17 17 Nicaraguan 17] border , Arce said the government has advised many to take additional security measures on their own . Hurte Sierd Zylstra and his wife , Jetsi Hendrika Coers , both 50 years old , were seized late on Saturday or early on Sunday from a teak tree plantation they manage by at least two heavily armed men who took the two off in their own car , leaving behind a ransom note demanding \$ 1.5 million . The plantation is owned by [18 18 Du tch 18] citizen Ebe Huizinga , who has since arrived in [19 19 Costa Rica 19] to deal with the matter .

1: 5551 Costa_Rica
2: 21148 Netherlands
3: 1170078 Nicaraguans
4: 219665 San_José,_Costa_Rica
5: 5551 Costa_Rica
6: 5551 Costa_Rica
7: 21148 Netherlands
8: 5551 Costa_Rica
9: 21362 Nicaragua
10: 21362 Nicaragua
11: 11867 Germany
12: 26748 Switzerland
13: 5551 Costa_Rica
14: 21362 Nicaragua
15: 21362 Nicaragua
16: 5551 Costa_Rica
17: 21362 Nicaragua
18: 21148 Netherlands
19: 5551 Costa_Rica
1: 5551 Costa_Rica
[pem_pos=0, final_scores=0.63387996,
final_local_scores=0.59923726, pem=-0.30652362,
similarity=23.639072]
2: 21148 Netherlands
[pem_pos=0, final_scores=0.4984843,
final_local_scores=0.51255655, pem=-0.7943452,
similarity=23.705196]
3: 1170078 Nicaraguans
[pem_pos=1, final_scores=0.1139044,
final_local_scores=0.14829314, pem=-0.66201633,
similarity=12.115223]
4: 219665 San_José,_Costa_Rica
[pem_pos=1, final_scores=0.55972195,
final_local_scores=0.5680813, pem=-2.9430156,
similarity=37.04697]
5: 5551 Costa_Rica
[pem_pos=0, final_scores=0.58525056,

```
final_local_scores=0.54611695, pem=-0.30652362,
similarity=22.053856]
6: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.6189167,
final_local_scores=0.5828922, pem=0.0, similarity=21.484379]
7: 21148 Netherlands
    [pem_pos=0, final_scores=0.57173324,
final_local_scores=0.5925699, pem=-0.7943452,
similarity=26.092953]
8: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.5915727,
final_local_scores=0.5530229, pem=-0.30652362,
similarity=22.259943]
9: 21362 Nicaragua
    [pem_pos=0, final_scores=0.545315,
final_local_scores=0.5015361, pem=0.0, similarity=19.05655]
10: 21362 Nicaragua
    [pem_pos=0, final_scores=0.49839288,
final_local_scores=0.45028073, pem=0.0, similarity=17.526987]
11: 11867 Germany
    [pem_pos=1, final_scores=0.6931151,
final_local_scores=0.7327449, pem=-1.0465823,
similarity=31.64776]
12: 26748 Switzerland
    [pem_pos=0, final_scores=0.68434817,
final_local_scores=0.71260333, pem=-0.1392413,
similarity=26.112434]
13: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.63030237,
final_local_scores=0.5953292, pem=-0.30652362,
similarity=23.522448]
14: 21362 Nicaragua
    [pem_pos=0, final_scores=0.49790877,
final_local_scores=0.4497519, pem=0.0, similarity=17.511208]
15: 21362 Nicaragua
    [pem_pos=0, final_scores=0.51784235,
final_local_scores=0.47152632, pem=0.0, similarity=18.161]
16: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.618069,
final_local_scores=0.58196616, pem=-0.30652362,
similarity=23.123667]
17: 21362 Nicaragua
    [pem_pos=0, final_scores=0.5249364,
final_local_scores=0.47927552, pem=0.0, similarity=18.39225]
18: 21148 Netherlands
    [pem_pos=0, final_scores=0.5521521,
final_local_scores=0.57118046, pem=-0.7943452,
similarity=25.45465]
19: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.50742173,
final_local_scores=0.46110082, pem=-0.30652362,
similarity=19.516804]
```

Sample: b"725_Gov't"

Gov't dodging extradition , [1 1 Colombian 1] official says . [2 2 BOGOTA 2] , [3 3 Colombia 3] 1996-08-28 A top judicial official and critic of President [4 4 Ernesto Samper 4] accused the government of indifference on Wednesday over efforts to lift [5 5 Colombia 5] 's five-year-old ban on extradition . " It would seem that the subject of extradition is unworthy of an opinion from the government , " Deputy Prosecutor-General Adolfo Salamanca said . Constitutional reforms were proposed on Tuesday by two senators , one of them a member of [6 6 Samper 6] 's own [7 7 Liberal Party 7] , aimed at lifting the ban on extradition introduced in 1991 . [8 8 U.S. 8] Ambassador Myles Frechette applauded the move , saying it could prompt the [9 9 Clinton 9] administration to remove [10 10 Colombia 10] from a list of outcast nations that have failed to cooperate in [11 11 U.S. 1 1] counternarcotics efforts . [12 12 Samper 12] -- who weathered a year-old crisis stemming from charges he financed his 1994 election campaign with drug money -- appeared less than enthusiastic , however . " Extradition is not on the government 's legislative agenda , " he told reporters on Tuesday . He added that he did not oppose the idea of opening a public debate over the issue . But he fell far short of endorsing the idea of putting [13 13 Colombian 13] drug lords onto U.S.-bound flights to serve stiff penalties in [14 14 American 14] prisons . Salamanca , who spoke at a meeting on kidnapping in [15 15 Colombia 15] , has said in the past that there was ample evidence to prove that [16 16 Samper 16] 's campaign received millions of dollars in contributions from the country 's top drug lords .

1: 5222 Colombia

2: 211271 Bogotá

3: 5222 Colombia

4: 974131 Ernesto_Samper

5: 5222 Colombia

6: 974131 Ernesto_Samper

7: 406013 Colombian_Liberal_Party

8: 3434750 United_States

9: 3356 Bill_Clinton

10: 5222 Colombia

11: 3434750 United_States

12: 974131 Ernesto_Samper

13: 5222 Colombia

14: 3434750 United_States

15: 5222 Colombia

16: 974131 Ernesto_Samper

1: 5222 Colombia

[pem_pos=0, final_scores=0.5442777,
final_local_scores=0.5236664, pem=-0.32036036,
similarity=21.459131]

2: 211271 Bogotá

```
[pem_pos=0, final_scores=0.6125453,
final_local_scores=0.6181953, pem=0.0, similarity=22.537893]
3: 5222 Colombia
    [pem_pos=0, final_scores=0.57482916,
final_local_scores=0.55703914, pem=-0.22328302,
similarity=21.92712]
4: 974131 Ernesto_Samper
    [pem_pos=0, final_scores=0.50022966,
final_local_scores=0.49108016, pem=0.0, similarity=18.744524]
5: 5222 Colombia
    [pem_pos=0, final_scores=0.6470155,
final_local_scores=0.63589174, pem=-0.22328302,
similarity=24.28024]
6: 974131 Ernesto_Samper
    [pem_pos=0, final_scores=0.38498896,
final_local_scores=0.37450367, pem=0.0, similarity=15.265649]
7: 406013 Colombian_Liberal_Party
    [pem_pos=12, final_scores=0.26073426,
final_local_scores=0.27038974, pem=-3.7501097,
similarity=32.55236]
8: 3434750 United_States
    [pem_pos=0, final_scores=0.5098663,
final_local_scores=0.53455347, pem=-0.14075686,
similarity=20.807314]
9: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.4900053,
final_local_scores=0.542406, pem=-1.920079,
similarity=30.71788]
10: 5222 Colombia
    [pem_pos=0, final_scores=0.46005172,
final_local_scores=0.4316622, pem=-0.22328302,
similarity=18.185621]
11: 3434750 United_States
    [pem_pos=0, final_scores=0.44908935,
final_local_scores=0.4681639, pem=-0.14075686,
similarity=18.826117]
12: 974131 Ernesto_Samper
    [pem_pos=0, final_scores=0.43760175,
final_local_scores=0.4319752, pem=0.0, similarity=16.980715]
13: 5222 Colombia
    [pem_pos=0, final_scores=0.543881,
final_local_scores=0.523233, pem=-0.32036036,
similarity=21.446198]
14: 3434750 United_States
    [pem_pos=0, final_scores=0.48274732,
final_local_scores=0.5049301, pem=-0.12994486,
similarity=19.864494]
15: 5222 Colombia
    [pem_pos=0, final_scores=0.5252419,
final_local_scores=0.5028726, pem=-0.22328302,
similarity=20.310682]
16: 974131 Ernesto_Samper
```

```
[pem_pos=0, final_scores=0.45093924,
final_local_scores=0.4465443, pem=0.0, similarity=17.415485]
```

Sample: b'726_Quake'

Quake shakes [1 1 Costa Rica 1] during [2 2 Hashimoto 2] visit . [3 3 SAN JOSE 3] , [4 4 Costa Rica 4] 1996-08-28 A moderate earthquake measuring 5.0 on the [5 5 Richter 5] scale shook [6 6 Costa Rica 6] on Wednesday during a visit by [7 7 Japanese 7] Prime Minister [8 8 Ryutaro Hashimoto 8] , but there were no reports of casualties or damage , officials said . The quake struck at 11.16 a.m. (1716 [9 9 GMT 9]) and was centred 10 miles (16 km) south of the port of Quepos , which is 90 miles (140 km) south of the capital [10 10 San Jose 10] , the Costa Rican Volcanic and Seismological Observatory said . The quake was felt for about seven seconds in most of the country but preliminary reports said no one was hurt , it added . The quake took place a few minutes before the end of a welcoming ceremony at Juan Santamaria airport for [11 11 Hashimoto 11] , who was starting a three-hour visit as part of a [12 12 Latin American 12] tour . [13 1 3 Hashimoto 13] , who arrived at 11 a.m. (1700 [14 14 GMT 14]) , showed no sign of having felt the quake , witnesses said .

1: 5551 Costa_Rica

2: 368549 Ryutaro_Hashimoto

3: 219665 San_José,_Costa_Rica

4: 5551 Costa_Rica

5: 19467352 Richter_magnitude_scale

6: 5551 Costa_Rica

7: 15573 Japan

8: 368549 Ryutaro_Hashimoto

9: 12701 Greenwich_Mean_Time

10: 219665 San_José,_Costa_Rica

11: 368549 Ryutaro_Hashimoto

12: 18524 Latin_America

13: 368549 Ryutaro_Hashimoto

14: 12701 Greenwich_Mean_Time

1: 5551 Costa_Rica

```
[pem_pos=0, final_scores=0.5051817,
final_local_scores=0.5664812, pem=-0.30652362,
similarity=22.661564]
```

2: 368549 Ryutaro_Hashimoto

```
[pem_pos=0, final_scores=0.9461587,
final_local_scores=1.0064323, pem=0.0, similarity=34.12366]
```

3: 219665 San_José,_Costa_Rica

```
[pem_pos=1, final_scores=0.57814723,
final_local_scores=0.62072307, pem=-2.9430156,
similarity=38.617905]
```

4: 5551 Costa_Rica

```
[pem_pos=0, final_scores=0.5071029,
final_local_scores=0.56857973, pem=-0.30652362,
similarity=22.72419]
```

5: 19467352 Richter_magnitude_scale

```
[pem_pos=0, final_scores=0.49414194,
final_local_scores=0.6091833, pem=-0.7852703,
similarity=26.539381]
6: 5551 Costa_Rica
    [pem_pos=0, final_scores=0.44306362,
final_local_scores=0.4986266, pem=-0.30652362,
similarity=20.636648]
7: 15573 Japan
    [pem_pos=0, final_scores=0.73557484,
final_local_scores=0.7723283, pem=-0.99076414,
similarity=32.52546]
8: 368549 Ryutaro_Hashimoto
    [pem_pos=0, final_scores=0.8552714,
final_local_scores=0.8925695, pem=0.0, similarity=30.725765]
9: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.6113991,
final_local_scores=0.72700924, pem=0.0, similarity=25.785118]
10: 219665 San_José,_Costa_Rica
    [pem_pos=1, final_scores=0.3663242,
final_local_scores=0.38933855, pem=-2.9430156,
similarity=31.712929]
11: 368549 Ryutaro_Hashimoto
    [pem_pos=0, final_scores=0.9173594,
final_local_scores=0.9749734, pem=0.0, similarity=33.184864]
12: 18524 Latin_America
    [pem_pos=0, final_scores=0.36565226,
final_local_scores=0.39774787, pem=-0.13736542,
similarity=16.706318]
13: 368549 Ryutaro_Hashimoto
    [pem_pos=0, final_scores=0.90930337,
final_local_scores=0.96617347, pem=0.0, similarity=32.922256]
14: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.6224045,
final_local_scores=0.7390309, pem=0.0, similarity=26.14387]
```

Sample: b'727_Barrier'

Barrier removed to [1 1 Brazil 1] [2 2 CVRD 2] sell-off . [3 3 BRASILIA 3] 1996-08-28 The [4 4 Brazilian 4] Senate Wednesday agreed to shelve a bill linking the privatization of mining conglomerate [5 5 Vale do Rio Doce 5] to congressional approval , officials said . Officials said the Senate vote removed all existing legislative hurdles in the way of [6 6 CVRD 6] 's sell-off . The motion was put forward by Sen . Jose Eduardo Dutra , who had drawn up the bill . The Senate vote also annulled a substitute version of Dutra 's bill which had sought to dedicate revenue from [7 7 Vale 7] 's privatization to regional infrastructure projects . -- William Schomberg , [8 8 Brasilia 8] newsroom 55-61-2230358

```
1: 3383 Brazil
2: 1873300 Vale_(mining_company)
3: 4752 Brasília
4: 3383 Brazil
```

```
5: 1873300 Vale_(mining_company)
6: 1873300 Vale_(mining_company)
7: 1873300 Vale_(mining_company)
8: 4752 Brasília
1: 3383 Brazil
    [pem_pos=0, final_scores=0.728723,
final_local_scores=0.75990087, pem=-0.21313342,
similarity=27.925722]
2: 1873300 Vale_(mining_company)
    [pem_pos=0, final_scores=0.5621396,
final_local_scores=0.53664917, pem=0.0, similarity=20.104395]
3: 4752 Brasília
    [pem_pos=0, final_scores=0.43125236,
final_local_scores=0.42156053, pem=0.0, similarity=16.66992]
4: 3383 Brazil
    [pem_pos=0, final_scores=0.79235786,
final_local_scores=0.8294123, pem=-0.2733083,
similarity=30.327324]
5: 1873300 Vale_(mining_company)
    [pem_pos=0, final_scores=0.5922978,
final_local_scores=0.5695925, pem=0.0, similarity=21.087488]
6: 1873300 Vale_(mining_company)
    [pem_pos=0, final_scores=0.60573894,
final_local_scores=0.5842749, pem=0.0, similarity=21.52564]
7: 1873300 Vale_(mining_company)
    [pem_pos=0, final_scores=0.52947205,
final_local_scores=0.5009649, pem=-0.8493678,
similarity=23.658499]
8: 4752 Brasília
    [pem_pos=0, final_scores=0.30207074,
final_local_scores=0.2804492, pem=0.0, similarity=12.458877]
```

Sample: b'728_RTRS'

RTRS - Arthur Yates year net A\$ 6.1 mln . [1 1 SYDNEY 1]
1996-08-29 Year to June 30 (million A\$ unless stated)
Operating profit 9.75 vs 5.79 Net profit 6.08 vs 3.98 Final
dividend (cents) 4.0 vs 4.0 Total dividend (cents) 6.0 vs
6.0 NOTE : Arthur Yates and Co ltd is a garden products
group . Sales 148.29 vs 133.82 Other income 1.90 vs 2.07 Shr
(cents) 8.63 vs 7.23 Dividend is 100 percent franked Pay
date Nov 25 Reg date Nov 11 Tax 3.67 vs 1.82 Interest 2.78 vs
2.69 Depreciation 3.25 vs 2.79 -- [2 2 Sydney 2] newsroom
61-2 9373-1800

```
1: 27862 Sydney
2: 27862 Sydney
1: 27862 Sydney
    [pem_pos=0, final_scores=0.68351763,
final_local_scores=0.61854774, pem=-0.1443489,
similarity=23.333405]
2: 27862 Sydney
    [pem_pos=0, final_scores=0.5115766,
final_local_scores=0.43072826, pem=-0.1443489,
```

```
similarity=17.728497]
```

Sample: b'729_Briton'

[1 1 Briton 1] held in [2 2 Thailand 2] over 4.4 kg heroin find . [3 3 BANGKOK 3] 1996-08-28 [4 4 Thai 4] airport police arrested a [5 5 British 5] bartender for allegedly attempting to board a flight for [6 6 Amsterdam 6] with nearly 4.4 kg (9.68 lb) of heroin in his luggage , police said on Wednesday . Police said James Lee Williams , 28 , was stopped at a [7 7 Bangkok 7] airport departure lounge on Monday after officials found the drug in a bag that Williams planned to carry onto the plane . Williams ' hometown was not immediately available . The maximum sentence for heroin trafficking is the death penalty , although it is normally commuted to life imprisonment .

1: 19097669 British_people

2: 30128 Thailand

3: 56615 Bangkok

4: 30128 Thailand

5: 31717 United_Kingdom

6: 844 Amsterdam

7: 56615 Bangkok

1: 19097669 British_people

[pem_pos=0, final_scores=0.1563255, final_local_scores=0.1792888, pem=-0.89919627, similarity=14.330017]

2: 30128 Thailand

[pem_pos=0, final_scores=0.6982804, final_local_scores=0.69455564, pem=-0.24029273, similarity=26.123386]

3: 56615 Bangkok

[pem_pos=0, final_scores=0.5062669, final_local_scores=0.47975034, pem=-0.0005963434, similarity=18.409662]

4: 30128 Thailand

[pem_pos=2, final_scores=0.6659081, final_local_scores=0.6591938, pem=-1.7504287, similarity=33.28047]

5: 31717 United_Kingdom

[pem_pos=0, final_scores=0.54187363, final_local_scores=0.516329, pem=-0.48936236, similarity=22.15923]

6: 844 Amsterdam

[pem_pos=0, final_scores=0.5982426, final_local_scores=0.6351608, pem=-0.0642803, similarity=23.393745]

7: 56615 Bangkok

[pem_pos=0, final_scores=0.5332597, final_local_scores=0.5092358, pem=-0.0005963434, similarity=19.289568]

Sample: b'730_PL0'

PLO Council calls for halt to contacts with [1 1 Israel 1] . [2 2 RAMALLAH 2] , [3 3 West Bank 3] 1996-08-28 The [4 4 Palestinian Legislative Council 4] on Wednesday called for a halt to contacts with [5 5 Israel 5] , just hours after President [6 6 Yasser Arafat 6] said the [1 Jewish 1] state had effectively declared war on the [7 7 Palestinians 7] by pursuing its hardline policies . A resolution released by the council called for " halting contacts with the [8 8 Israeli 8] side and leaving the mechanism to carry out this to [9 9 Palestinian 9] President [10 10 Yasser Arafat 10] " . The council was meeting in [11 11 Ramallah 11] to discuss [12 12 Israel 12] 's new policy of [2 Jewish 2] settlement expansion and its uncompromising line on [13 13 Jerusalem 13] since Prime Minister [14 14 Benjamin Netanyahu 14] took office in June . Council resolutions are not necessarily binding . [15 15 Arafat 15] had earlier blasted [16 16 Israel 16] saying its policies amounted to a declaration of war against the [17 17 Palestinian 17] people . He also called for the first general strike in two years in the [18 18 West Bank 18] and [19 19 Gaza 19] on Thursday . " What happened concerning continuous violations and crimes from this new [20 20 Israeli 20] leadership means they are declaring a state of war against the [21 21 Palestinian 21] people , " [22 22 Arafat 22] told the council . Council speaker Ahmed Korei said the decision was part of a comprehensive plan to confront [23 23 Israeli 23] settlement policy , land confiscation and what he termed other violations of the Israeli-PLO peace deals .

1: 9282173 Israel
2: 44132 Ramallah
3: 96665 Palestinian_territories
4: 40067464 Palestinian_Legislative_Council
5: 9282173 Israel
6: 34289 Yasser_Arafat
7: 241405 State_of_Palestine
8: 9282173 Israel
9: 241405 State_of_Palestine
10: 34289 Yasser_Arafat
11: 44132 Ramallah
12: 9282173 Israel
13: 16043 Jerusalem
14: 56472 Benjamin_Netanyahu
15: 34289 Yasser_Arafat
16: 9282173 Israel
17: 241405 State_of_Palestine
18: 96665 Palestinian_territories
19: 55387 Gaza
20: 9282173 Israel
21: 241405 State_of_Palestine
22: 34289 Yasser_Arafat
23: 9282173 Israel
1: 9282173 Israel
[pem_pos=0, final_scores=0.6155905,

```
final_local_scores=0.5720903, pem=-0.21279564,
similarity=22.319244]
2: 44132 Ramallah
    [pem_pos=0, final_scores=0.4606635,
final_local_scores=0.4253223, pem=0.0, similarity=16.782179]
3: 96665 Palestinian_territories
    [pem_pos=1, final_scores=0.31237084,
final_local_scores=0.26183772, pem=-1.929408,
similarity=22.395893]
4: 40067464 Palestinian_Legislative_Council
    [pem_pos=0, final_scores=0.4151492,
final_local_scores=0.42373043, pem=0.0, similarity=16.734674]
5: 9282173 Israel
    [pem_pos=0, final_scores=0.61360884,
final_local_scores=0.5699256, pem=-0.21279564,
similarity=22.254644]
6: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.5959993,
final_local_scores=0.5678706, pem=0.0, similarity=21.036104]
7: 241405 State_of_Palestine
    [pem_pos=5, final_scores=0.46537858,
final_local_scores=0.44025517, pem=-1.7909594,
similarity=26.967321]
8: 9282173 Israel
    [pem_pos=0, final_scores=0.57421637,
final_local_scores=0.52689534, pem=-0.34790626,
similarity=21.705288]
9: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.7193962,
final_local_scores=0.71773076, pem=-2.8508017,
similarity=41.011333]
10: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.6432701,
final_local_scores=0.6195067, pem=0.0, similarity=22.577028]
11: 44132 Ramallah
    [pem_pos=0, final_scores=0.4578598,
final_local_scores=0.4222597, pem=0.0, similarity=16.690784]
12: 9282173 Israel
    [pem_pos=0, final_scores=0.58895475,
final_local_scores=0.5429948, pem=-0.21279564,
similarity=21.450975]
13: 16043 Jerusalem
    [pem_pos=0, final_scores=0.49760568,
final_local_scores=0.47636718, pem=-0.13239029,
similarity=19.02542]
14: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.63544214,
final_local_scores=0.63600284, pem=0.0, similarity=23.069305]
15: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.58778435,
final_local_scores=0.56281596, pem=0.0, similarity=20.885263]
16: 9282173 Israel
```

```
[pem_pos=0, final_scores=0.63548034,
final_local_scores=0.59381694, pem=-0.21279564,
similarity=22.96761]
17: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.54902107,
final_local_scores=0.53162193, pem=-2.8508017,
similarity=35.457474]
18: 96665 Palestinian_territories
    [pem_pos=1, final_scores=0.34883708,
final_local_scores=0.30167162, pem=-1.929408,
similarity=23.584616]
19: 55387 Gaza
    [pem_pos=0, final_scores=0.23992212,
final_local_scores=0.1862579, pem=-0.30319342,
similarity=11.296833]
20: 9282173 Israel
    [pem_pos=0, final_scores=0.55631393,
final_local_scores=0.50733966, pem=-0.34790626,
similarity=21.121708]
21: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.48235506,
final_local_scores=0.45879936, pem=-2.8508017,
similarity=33.284306]
22: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.63308436,
final_local_scores=0.6122994, pem=0.0, similarity=22.361948]
23: 9282173 Israel
    [pem_pos=0, final_scores=0.51184744,
final_local_scores=0.45876676, pem=-0.34790626,
similarity=19.672197]
1: 25955086 Jews
    [pem_pos=1, final_scores=0.066112176,
final_local_scores=0.035602063, pem=-0.80476934,
similarity=9.52861]
2: 25955086 Jews
    [pem_pos=1, final_scores=0.08417876,
final_local_scores=0.05533707, pem=-0.80476934,
similarity=10.117542]
```

Sample: b'731_Iraqi'

[1 Iraqi Kurd 1] group says agrees new ceasefire . [1 1 ANKA RA 1] 1996-08-28 An [2 2 Iraqi Kurdish 2] group on Wednesday said it had agreed a new U.S.-brokered ceasefire with a rival faction after a previous accord was shattered by sporadic fighting between the groups in recent days . " The [3 3 Patriotic Union of Kurdistan 3] ([4 4 PUK 4]) leadership declares its endorsement for a ceasefire arrangement with the [5 5 KDP 5] ([6 6 Kurdistan Democratic Party 6]) to take effect as of 8:00 a.m. on August 28 , " the [7 7 PUK 7] said in a statement . The [8 8 PUK 8] said the ceasefire was agreed after talks between [9 9 U.S. 9] Assistant Secretary for Near East Affairs [10 10 Robert Pelletreau 10] and [11 11

PUK 11] leader [12 12 Jalal Talabani 12] . The [13 13 KDP 13] , led by [14 14 Massoud Barzani 14] , had said a previous ceasefire negotiated by [15 15 Pelletreau 15] last Friday was broken by the [16 16 PUK 16] . [17 17 Talabani 17] has agreed to take part in talks in [18 18 London 18] on reaching a comprehensive settlement for the PUK-KDP conflict , the [19 19 PUK 19] statement said . It said the [20 20 KDP 20] was responsible for breaking the previous ceasefire by refusing to endorse it publicly .

1: 802 Ankara
2: 679693 Iraqi_Kurdistan
3: 70789 Patriotic_Union_of_Kurdistan
4: 70789 Patriotic_Union_of_Kurdistan
5: 25659771 Kurdistan_Democratic_Party
6: 25659771 Kurdistan_Democratic_Party
7: 70789 Patriotic_Union_of_Kurdistan
8: 70789 Patriotic_Union_of_Kurdistan
9: 3434750 United_States
10: 18299389 Robert_Pelletreau
11: 70789 Patriotic_Union_of_Kurdistan
12: 440244 Jalal_Talabani
13: 25659771 Kurdistan_Democratic_Party
14: 8122677 Masoud_Barzani
15: 18299389 Robert_Pelletreau
16: 70789 Patriotic_Union_of_Kurdistan
17: 440244 Jalal_Talabani
18: 17867 London
19: 70789 Patriotic_Union_of_Kurdistan
20: 25659771 Kurdistan_Democratic_Party
1: 802 Ankara
[pem_pos=0, final_scores=0.3715163,
final_local_scores=0.41451073, pem=0.0, similarity=16.45954]
2: 679693 Iraqi_Kurdistan
[pem_pos=0, final_scores=0.4026683,
final_local_scores=0.36773401, pem=0.0, similarity=15.06363]
3: 70789 Patriotic_Union_of_Kurdistan
[pem_pos=0, final_scores=0.6591386,
final_local_scores=0.6241976, pem=0.0, similarity=22.717014]
4: 70789 Patriotic_Union_of_Kurdistan
[pem_pos=0, final_scores=0.54085416,
final_local_scores=0.49498975, pem=0.0, similarity=18.861195]
5: 25659771 Kurdistan_Democratic_Party
[pem_pos=0, final_scores=0.6759818,
final_local_scores=0.6477125, pem=0.0, similarity=23.418747]
6: 25659771 Kurdistan_Democratic_Party
[pem_pos=0, final_scores=0.58611816,
final_local_scores=0.5495501, pem=0.0, similarity=20.489384]
7: 70789 Patriotic_Union_of_Kurdistan
[pem_pos=0, final_scores=0.5813623,
final_local_scores=0.5392388, pem=0.0, similarity=20.181675]
8: 70789 Patriotic_Union_of_Kurdistan
[pem_pos=0, final_scores=0.54749435,

```
final_local_scores=0.5022432, pem=0.0, similarity=19.077652]
9: 3434750 United_States
    [pem_pos=0, final_scores=0.44737715,
final_local_scores=0.5308811, pem=-0.14075686,
similarity=20.697723]
10: 18299389 Robert_Pelletreau
    [pem_pos=0, final_scores=0.43415552,
final_local_scores=0.5799083, pem=0.0, similarity=21.395334]
11: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.5491665,
final_local_scores=0.5040698, pem=0.0, similarity=19.132162]
12: 440244 Jalal_Talabani
    [pem_pos=0, final_scores=0.5014464,
final_local_scores=0.49005944, pem=0.0, similarity=18.714064]
13: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.50696653,
final_local_scores=0.46308905, pem=0.0, similarity=17.909214]
14: 8122677 Masoud_Barzani
    [pem_pos=0, final_scores=0.59377927,
final_local_scores=0.5967143, pem=0.0, similarity=21.896858]
15: 18299389 Robert_Pelletreau
    [pem_pos=0, final_scores=0.37098664,
final_local_scores=0.5109059, pem=0.0, similarity=19.336164]
16: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.566236,
final_local_scores=0.5227156, pem=0.0, similarity=19.68859]
17: 440244 Jalal_Talabani
    [pem_pos=0, final_scores=0.41324413,
final_local_scores=0.3937118, pem=0.0, similarity=15.838859]
18: 17867 London
    [pem_pos=0, final_scores=0.4459561,
final_local_scores=0.5508191, pem=-0.09086737,
similarity=21.021404]
19: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.5770454,
final_local_scores=0.5345232, pem=0.0, similarity=20.04095]
20: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.5893008,
final_local_scores=0.55302674, pem=0.0, similarity=20.593134]
1: 679693 Iraqi_Kurdistan
    [pem_pos=0, final_scores=0.110555604,
final_local_scores=0.048645303, pem=0.0, similarity=5.5413857]
```

Sample: b'732_Kurd'

[1 1 Kurd 1] group says [2 2 Iraqi 2] troops massing near north . [3 3 ANKARA 3] 1996-08-28 An [4 4 Iraqi Kurdish 4] group on Wednesday said [5 5 Iraq 5] was massing troops near [6 6 Kurdish 6] regions in the north , where a U.S.-led allied air force protects the local population against attacks from [7 7 Baghdad 7] . " The [8 8 Iraqi 8] regime has started threatening the [9 9 Kurdish 9] population by massing troops in preparation to attack [10 10 Kurdish 10] towns and

population centres , " the [11 11 Patriotic Union of Kurdistan 11] ([12 12 PUK 12]) said in a statement . The [13 13 PUK 13]] said it had received confirmed reports that [14 14 Iraqi 14] troops , supported by tanks , artillery and armoured vehicles , have already penetrated some [15 15 Kurdish 15] areas . It said the military presence reflects cooperation between President [16 16 Saddam Hussein 16] and the [17 17 PUK 17] 's rival , the [18 18 Kurdistan Democratic Party 18] ([19 19 KDP 19]) . The [20 20 PUK 20] statement follows [21 21 KDP 21] assertions that the [22 22 PUK 22] is receiving military support from [23 23 Iran 23] . Hostilities between the two warring [24 24 Iraqi Kurdish 24] factions have continued in the last few days despite a U.S.-brokered ceasefire last Friday . The [25 25 PUK 25] called on the [26 26 United Nations 26] and allied forces to halt the [27 27 Iraqi 27] aggression . [28 28 U.S. 28] , [29 29 French 29] and [30 30 British 30] aircraft have safeguarded the [31 31 Iraqi Kurdish 31] population against aggression from [32 32 Baghdad 32] since shortly after the [33 33 Gulf War 33] in 1991 . The allied force , known as [34 34 Operation Provide Comfort 34] , is based in southern Turkey .

1: 17068 Kurdish_people
2: 7515928 Iraq
3: 802 Ankara
4: 679693 Iraqi_Kurdistan
5: 7515928 Iraq
6: 80777 Kurdistan
7: 4492 Baghdad
8: 7515928 Iraq
9: 80777 Kurdistan
10: 80777 Kurdistan
11: 70789 Patriotic_Union_of_Kurdistan
12: 70789 Patriotic_Union_of_Kurdistan
13: 70789 Patriotic_Union_of_Kurdistan
14: 7515928 Iraq
15: 80777 Kurdistan
16: 29490 Saddam_Hussein
17: 70789 Patriotic_Union_of_Kurdistan
18: 25659771 Kurdistan_Democratic_Party
19: 25659771 Kurdistan_Democratic_Party
20: 70789 Patriotic_Union_of_Kurdistan
21: 25659771 Kurdistan_Democratic_Party
22: 70789 Patriotic_Union_of_Kurdistan
23: 14653 Iran
24: 679693 Iraqi_Kurdistan
25: 70789 Patriotic_Union_of_Kurdistan
26: 31769 United_Nations
27: 7515928 Iraq
28: 3434750 United_States
29: 5843419 France
30: 31717 United_Kingdom
31: 679693 Iraqi_Kurdistan

32: 4492 Baghdad
33: 182000 Gulf_War
34: 252171 Operation_Provide_Comfort
1: 17068 Kurdish_people
[pem_pos=0, final_scores=0.25078923,
final_local_scores=0.19794136, pem=0.0, similarity=9.996678]
2: 7515928 Iraq
[pem_pos=0, final_scores=0.63127613,
final_local_scores=0.58758473, pem=0.0, similarity=21.624413]
3: 802 Ankara
[pem_pos=0, final_scores=0.4050532,
final_local_scores=0.4180649, pem=0.0, similarity=16.565603]
4: 679693 Iraqi_Kurdistan
[pem_pos=0, final_scores=0.41159087,
final_local_scores=0.35881978, pem=0.0, similarity=14.797611]
5: 7515928 Iraq
[pem_pos=0, final_scores=0.71306723,
final_local_scores=0.67529935, pem=-0.25379938,
similarity=25.622192]
6: 80777 Kurdistan
[pem_pos=9, final_scores=0.37468076,
final_local_scores=0.32544017, pem=-2.8001814,
similarity=29.029318]
7: 4492 Baghdad
[pem_pos=0, final_scores=0.31775415,
final_local_scores=0.2644508, pem=0.0, similarity=11.981452]
8: 7515928 Iraq
[pem_pos=0, final_scores=0.60547537,
final_local_scores=0.5577716, pem=-0.48469657,
similarity=23.370586]
9: 80777 Kurdistan
[pem_pos=9, final_scores=0.3636542,
final_local_scores=0.31339538, pem=-2.8001814,
similarity=28.669878]
10: 80777 Kurdistan
[pem_pos=9, final_scores=0.322506,
final_local_scores=0.26844722, pem=-2.8001814,
similarity=27.328535]
11: 70789 Patriotic_Union_of_Kurdistan
[pem_pos=0, final_scores=0.6273915,
final_local_scores=0.62770927, pem=0.0, similarity=22.82181]
12: 70789 Patriotic_Union_of_Kurdistan
[pem_pos=0, final_scores=0.6009732,
final_local_scores=0.59885126, pem=0.0, similarity=21.960629]
13: 70789 Patriotic_Union_of_Kurdistan
[pem_pos=0, final_scores=0.53946835,
final_local_scores=0.5316666, pem=0.0, similarity=19.955704]
14: 7515928 Iraq
[pem_pos=0, final_scores=0.6693149,
final_local_scores=0.6291363, pem=0.0, similarity=22.864395]
15: 80777 Kurdistan
[pem_pos=9, final_scores=0.43085557,

```
final_local_scores=0.38680267, pem=-2.8001814,
similarity=30.860498]
16: 29490 Saddam_Hussein
    [pem_pos=0, final_scores=0.50137144,
final_local_scores=0.48172176, pem=0.0, similarity=18.465252]
17: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.61538523,
final_local_scores=0.61459416, pem=0.0, similarity=22.43043]
18: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.61637646,
final_local_scores=0.61074597, pem=0.0, similarity=22.31559]
19: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.58197683,
final_local_scores=0.5731696, pem=0.0, similarity=21.194237]
20: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.5285582,
final_local_scores=0.51974887, pem=0.0, similarity=19.600056]
21: 25659771 Kurdistan_Democratic_Party
    [pem_pos=0, final_scores=0.5144182,
final_local_scores=0.49937207, pem=0.0, similarity=18.991972]
22: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.56863314,
final_local_scores=0.56352466, pem=0.0, similarity=20.906412]
23: 14653 Iran
    [pem_pos=0, final_scores=0.61795574,
final_local_scores=0.6917955, pem=-0.13856095,
similarity=25.487787]
24: 679693 Iraqi_Kurdistan
    [pem_pos=0, final_scores=0.464377,
final_local_scores=0.4164806, pem=0.0, similarity=16.518324]
25: 70789 Patriotic_Union_of_Kurdistan
    [pem_pos=0, final_scores=0.56616217,
final_local_scores=0.5608255, pem=0.0, similarity=20.825865]
26: 31769 United_Nations
    [pem_pos=0, final_scores=0.45039797,
final_local_scores=0.47917432, pem=0.0, similarity=18.38923]
27: 7515928 Iraq
    [pem_pos=0, final_scores=0.69930553,
final_local_scores=0.6618965, pem=0.0, similarity=23.842026]
28: 3434750 United_States
    [pem_pos=0, final_scores=0.4305939,
final_local_scores=0.4678771, pem=-0.14075686,
similarity=18.817554]
29: 5843419 France
    [pem_pos=0, final_scores=0.5175436,
final_local_scores=0.57104194, pem=-0.80223274,
similarity=25.49341]
30: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.47947013,
final_local_scores=0.5129558, pem=-0.48936236,
similarity=22.058565]
31: 679693 Iraqi_Kurdistan
```

```
[pem_pos=0, final_scores=0.4413604,
final_local_scores=0.39133847, pem=0.0, similarity=15.768034]
32: 4492 Baghdad
[pem_pos=0, final_scores=0.4040696,
final_local_scores=0.3587373, pem=0.0, similarity=14.795149]
33: 182000 Gulf_War
[pem_pos=0, final_scores=0.37636292,
final_local_scores=0.36823648, pem=0.0, similarity=15.078625]
34: 252171 Operation_Provide_Comfort
[pem_pos=0, final_scores=0.3080657,
final_local_scores=0.35683966, pem=0.0, similarity=14.73852]
```

Sample: b'733_Iraq'

[1 1 Iraq 1] says hijackers were not diplomats . [2 2 BAGHDAD 2] 1996-08-28 [3 3 Iraq 3] on Wednesday said the hijackers of a [4 4 Sudanese 4] airliner were not [5 5 Iraqi 5] diplomats and added that " noble [6 6 Iraqis 6] " would never contemplate such an action . The official [7 7 Iraqi News Agency 7] (INA) quoted [8 8 Iraq 8] 's ambassador in [9 9 Khartoum 9] as saying that [10 10 Iraq 10] 's embassy in the [11 11 Sudanese 11] capital had nothing to do with the Monday night hijacking . [12 12 Iraq 12] 's ambassador in [13 13 Khartoum 13] denounced the hijacking and described it as a terrorist act which had nothing to do " with the morals and values of noble [14 14 Iraqis 14] , " INA said . Ambassador Abdulsamad Hameed Ali told INA there was only one diplomat among the 199 passengers and crew on the [15 15 Sudan Airways 15] Airbus . " He was not involved ... on the contrary he was harassed by the elements which carried out the hijacking , " he said . INA did not say the hijackers were [16 16 Iraqis 16] . The hijack started when the flight left [17 17 Khartoum 17] for [18 18 Amman 18] on Monday night . The hijackers told the crew they had grenades and other explosives and threatened to blow up the plane if they were not taken to [19 19 London 19] . The airliner refuelled at [20 20 Larnaca 20] , [21 21 Cyprus 21] and landed at [22 22 London 22] 's [23 23 Stansted 23] airport in the early hours of Tuesday . Seven [24 24 Iraqi 24] suspected hijackers surrendered and [25 25 British 25] police said they had apparently asked for political asylum . Several had brought their families along , including children .

1: 7515928 Iraq
2: 4492 Baghdad
3: 7515928 Iraq
4: 27421 Sudan
5: 7515928 Iraq
6: 7515928 Iraq
7: 8550616 National_Iraqi_News_Agency
8: 7515928 Iraq
9: 17320 Khartoum
10: 7515928 Iraq
11: 27421 Sudan
12: 7515928 Iraq

13: 17320 Khartoum
14: 7515928 Iraq
15: 383994 Sudan_Airways
16: 7515928 Iraq
17: 17320 Khartoum
18: 45007 Amman
19: 17867 London
20: 287852 Larnaca
21: 5593 Cyprus
22: 17867 London
23: 65456 London_Stansted_Airport
24: 7515928 Iraq
25: 31717 United_Kingdom
1: 7515928 Iraq
[pem_pos=0, final_scores=0.6355261,
final_local_scores=0.58438474, pem=-0.25379938,
similarity=22.90912]
2: 4492 Baghdad
[pem_pos=0, final_scores=0.45877355,
final_local_scores=0.4199403, pem=0.0, similarity=16.621569]
3: 7515928 Iraq
[pem_pos=0, final_scores=0.6119277,
final_local_scores=0.558607, pem=-0.25379938,
similarity=22.13986]
4: 27421 Sudan
[pem_pos=0, final_scores=0.41878968,
final_local_scores=0.37807637, pem=-0.19893567,
similarity=16.45411]
5: 7515928 Iraq
[pem_pos=0, final_scores=0.5966232,
final_local_scores=0.54188913, pem=-0.48469657,
similarity=22.896622]
6: 7515928 Iraq
[pem_pos=1, final_scores=0.5077139,
final_local_scores=0.44476926, pem=-0.8129937,
similarity=21.7837]
7: 8550616 National_Iraqi_News_Agency
[pem_pos=0, final_scores=0.39496952,
final_local_scores=0.588533, pem=0.0, similarity=21.652712]
8: 7515928 Iraq
[pem_pos=0, final_scores=0.64973205,
final_local_scores=0.5999025, pem=-0.25379938,
similarity=23.372202]
9: 17320 Khartoum
[pem_pos=0, final_scores=0.3686509,
final_local_scores=0.32586348, pem=-0.03966028,
similarity=14.029808]
10: 7515928 Iraq
[pem_pos=0, final_scores=0.6477037,
final_local_scores=0.5976869, pem=-0.25379938,
similarity=23.306082]
11: 27421 Sudan

```
[pem_pos=0, final_scores=0.4313708,
final_local_scores=0.39181924, pem=-0.19893567,
similarity=16.864223]
12: 7515928 Iraq
    [pem_pos=0, final_scores=0.6296057,
final_local_scores=0.5779175, pem=-0.25379938,
similarity=22.716125]
13: 17320 Khartoum
    [pem_pos=0, final_scores=0.3393404,
final_local_scores=0.29384616, pem=-0.03966028,
similarity=13.0743475]
14: 7515928 Iraq
    [pem_pos=1, final_scores=0.46416563,
final_local_scores=0.3971994, pem=-0.8129937,
similarity=20.364122]
15: 383994 Sudan_Airways
    [pem_pos=0, final_scores=0.5808635,
final_local_scores=0.6784238, pem=0.0, similarity=24.335234]
16: 7515928 Iraq
    [pem_pos=1, final_scores=0.51108193,
final_local_scores=0.4484483, pem=-0.8129937,
similarity=21.893492]
17: 17320 Khartoum
    [pem_pos=0, final_scores=0.37474906,
final_local_scores=0.33252478, pem=-0.03966028,
similarity=14.228595]
18: 45007 Amman
    [pem_pos=0, final_scores=0.41364843,
final_local_scores=0.3833791, pem=0.0, similarity=15.53051]
19: 17867 London
    [pem_pos=0, final_scores=0.4922772,
final_local_scores=0.51047736, pem=-0.09086737,
similarity=19.817526]
20: 287852 Larnaca
    [pem_pos=0, final_scores=0.38878423,
final_local_scores=0.4204594, pem=0.0, similarity=16.63706]
21: 5593 Cyprus
    [pem_pos=0, final_scores=0.61242473,
final_local_scores=0.6064498, pem=-0.3752463,
similarity=24.228031]
22: 17867 London
    [pem_pos=0, final_scores=0.5496995,
final_local_scores=0.5732025, pem=-0.09086737,
similarity=21.68937]
23: 65456 London_Stansted_Airport
    [pem_pos=0, final_scores=0.3799597,
final_local_scores=0.45772612, pem=-0.2670293,
similarity=19.20132]
24: 7515928 Iraq
    [pem_pos=0, final_scores=0.5922949,
final_local_scores=0.5371611, pem=-0.48469657,
similarity=22.755527]
```

25: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5082784,
final_local_scores=0.51186275, pem=-0.48936236,
similarity=22.025948]

Sample: b'734_Dole'
[1 1 Dole 1] blasts [2 2 Clinton 2] for ignoring teen drug use . Judith Crosson [3 3 VENTURA 3] , [4 4 Calif. 4] 1996-08-28 [5 5 Republican 5] presidential candidate [6 6 Bob Dole 6] Wednesday accused the [7 7 Clinton 7] administration of ignoring drug use among teenagers and said if elected he would use the [8 8 National Guard 8] to stop drugs from entering the [9 9 United States 9] . " He 'll probably mention his war on drugs , which he 's going to start like everything else -- next year . It 's too late , Mr. President , " [10 10 Dole 10] told an outdoor crowd of several hundred at a private religious school . He also commented briefly on published reports that the administration was planning to announce a plan to lower capital gains taxes for home sales . " Welcome to the club . We 've had it out there for weeks and weeks and weeks , " [11 11 Dole 11] said . [12 12 Dole 12] said former first lady [13 13 Nancy Reagan 13] was laughed at with her " just say no " anti-drug message . " But it worked , " [14 14 Dole 14] said . Meanwhile , in [15 15 Los Angeles 15] , [16 16 Dole 16] 's running mate , [17 17 Jack Kemp 17] , campaigned aggressively for the black vote in an area that was the flashpoint of the 1992 [18 18 Los Angeles 18] riots . [19 19 K emp 19] told a crowd of about 300 [20 20 African Americans 20] in south central [21 21 Los Angeles 21] , " Keep your eyes open , keep your ears open , keep your heart open . I want to tell you with all my heart that we want to win your vote . " In [22 22 Dole 22] 's address to a group that was largely white , the presidential nominee likened the stream of illegal drugs into the [23 23 United States 23] to missiles aimed at [24 24 American 24] children and promised to appoint federal judges who would be tough on illegal drug use . " They 're aiming millions and millions of missiles right at these young people , whether it 's a needle , whether it 's a cigarette , whatever the delivery system is -- it 's poison and it 's got to stop in [25 25 America 25] . " he said . [26 26 Dole 26] said 70 percent of the cocaine that entered the [27 27 United States 27] and 40 percent of the marijuana came from [28 28 Mexico 28] . " We 've got an international problem and I 'm prepared to use our military might . We want to stop drugs at the border , " he said . [29 29 Dole 29] 's remarks prompted questions about whether he was seeking a ban on cigarettes . " I did n't say anything about cigarettes . I was talking about drugs . I said you should n't smoke either . That 's all I said , " he replied as he was shaking hands with well-wishers . When asked specifically if he was suggesting a ban on cigarettes , [30 30 Dole 30] replied : "

Oh no . Come on , you know better than that . " [31 31 Dole 31] campaign aides said the candidate was telling young people not to smoke . [32 32 Dole 32] also said he opposed California Proposition 215 which , if approved by voters , would allow the cultivation of marijuana plants for medicinal uses . [33 33 Dole 33] said the initiative would allow marijuana to be used for anything from a headache to an ingrown toenail . In an effort to paint the drug issue in non-political terms , [34 34 Dole 34] said three times during his 20-minute address that illegal drug use was neither a [35 35 Democratic 35] nor a [36 36 Republican 36] issue but one that involves all people . The anti-drug message is a theme [37 37 Dole 37] feels has strong voter appeal . On Sunday near [38 38 Chicago 38] he accused President [39 39 Bill Clinton 39] of " raising the white flag " in the war on drugs . A recent survey showed that illegal drug use among 12-17 year-olds had doubled in the past four years . [40 40 Dole 40] was flanked by several [41 41 California 41] [42 42 Republican 42] politicians including Gov . [43 43 Pete Wilson 43] , who said local and state governments cannot fight illegal drugs alone . " We need all the help we can get . We need to get the kind of help we used to get when [44 44 Ronald Reagan 44] and [45 45 George Bush 45] were in the [46 46 White House 46] , " [47 47 Wilson 47] said .

1: 55846 Bob_Dole
2: 3356 Bill_Clinton
3: 108340 Ventura,_California
4: 5407 California
5: 32070 Republican_Party_(United_States)
6: 55846 Bob_Dole
7: 3356 Bill_Clinton
8: 37988 National_Guard_of_the_United_States
9: 3434750 United_States
10: 55846 Bob_Dole
11: 55846 Bob_Dole
12: 55846 Bob_Dole
13: 21181 Nancy_Reagan
14: 55846 Bob_Dole
15: 18110 Los_Angeles
16: 55846 Bob_Dole
17: 16200 Jack_Kemp
18: 18110 Los_Angeles
19: 16200 Jack_Kemp
20: 2154 African_American
21: 18110 Los_Angeles
22: 55846 Bob_Dole
23: 3434750 United_States
24: 3434750 United_States
25: 3434750 United_States
26: 55846 Bob_Dole
27: 3434750 United_States
28: 3966054 Mexico

```
29: 55846 Bob_Dole
30: 55846 Bob_Dole
31: 55846 Bob_Dole
32: 55846 Bob_Dole
33: 55846 Bob_Dole
34: 55846 Bob_Dole
35: 5043544 Democratic_Party_(United_States)
36: 32070 Republican_Party_(United_States)
37: 55846 Bob_Dole
38: 6886 Chicago
39: 3356 Bill_Clinton
40: 55846 Bob_Dole
41: 5407 California
42: 32070 Republican_Party_(United_States)
43: 343170 Pete_Wilson
44: 25433 Ronald_Reagan
45: 11955 George_H._W._Bush
46: 33057 White_House
47: 343170 Pete_Wilson
1: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.8437031,
final_local_scores=0.81511205, pem=0.0, similarity=28.414282]
2: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.877276,
final_local_scores=0.87926275, pem=0.0, similarity=30.328667]
3: 108340 Ventura,_California
    [pem_pos=0, final_scores=0.3540551,
final_local_scores=0.4084345, pem=-0.75008714,
similarity=20.357304]
4: 5407 California
    [pem_pos=0, final_scores=0.54472524,
final_local_scores=0.58786756, pem=0.0, similarity=21.632853]
5: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.43792707,
final_local_scores=0.40732938, pem=-0.1310191,
similarity=16.957737]
6: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.7819391,
final_local_scores=0.74692106, pem=0.0, similarity=26.379328]
7: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.8308615,
final_local_scores=0.82856196, pem=0.0, similarity=28.815655]
8: 37988 National_Guard_of_the_United_States
    [pem_pos=0, final_scores=0.39865226,
final_local_scores=0.44952863, pem=-0.3472341,
similarity=19.392859]
9: 3434750 United_States
    [pem_pos=0, final_scores=0.4821849,
final_local_scores=0.51100284, pem=-0.073457584,
similarity=19.738531]
10: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.8554246,
```

```
final_local_scores=0.827916, pem=0.0, similarity=28.79638]
11: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.9270198,
final_local_scores=0.90612286, pem=0.0, similarity=31.130226]
12: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.9416812,
final_local_scores=0.9221383, pem=0.0, similarity=31.60816]
13: 21181 Nancy_Reagan
    [pem_pos=0, final_scores=0.36182135,
final_local_scores=0.3624962, pem=0.0, similarity=14.907323]
14: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.9085936,
final_local_scores=0.8859951, pem=0.0, similarity=30.529573]
15: 18110 Los_Angeles
    [pem_pos=0, final_scores=0.60841674,
final_local_scores=0.6573849, pem=-0.100652285,
similarity=24.254751]
16: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.8767676,
final_local_scores=0.8512301, pem=0.0, similarity=29.492117]
17: 16200 Jack_Kemp
    [pem_pos=0, final_scores=0.37034088,
final_local_scores=0.3555222, pem=0.0, similarity=14.699204]
18: 18110 Los_Angeles
    [pem_pos=0, final_scores=0.51417834,
final_local_scores=0.5544437, pem=-0.100652285,
similarity=21.182781]
19: 16200 Jack_Kemp
    [pem_pos=0, final_scores=0.73763865,
final_local_scores=0.75926, pem=0.0, similarity=26.747545]
20: 2154 African_American
    [pem_pos=0, final_scores=0.39892578,
final_local_scores=0.41713953, pem=0.0, similarity=16.537989]
21: 18110 Los_Angeles
    [pem_pos=0, final_scores=0.47124654,
final_local_scores=0.5075473, pem=-0.100652285,
similarity=19.7833]
22: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.8983148,
final_local_scores=0.874767, pem=0.0, similarity=30.194504]
23: 3434750 United_States
    [pem_pos=0, final_scores=0.4739973,
final_local_scores=0.50205916, pem=-0.073457584,
similarity=19.471634]
24: 3434750 United_States
    [pem_pos=0, final_scores=0.5124545,
final_local_scores=0.5440678, pem=-0.12994486,
similarity=21.032442]
25: 3434750 United_States
    [pem_pos=0, final_scores=0.47957867,
final_local_scores=0.50815594, pem=-0.78483206,
similarity=23.522139]
```

```
26: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.9951474,
final_local_scores=0.980542, pem=0.0, similarity=33.351044]
27: 3434750 United_States
    [pem_pos=0, final_scores=0.4603837,
final_local_scores=0.4871884, pem=-0.073457584,
similarity=19.02786]
28: 3966054 Mexico
    [pem_pos=0, final_scores=0.50873375,
final_local_scores=0.6059712, pem=-0.18258883,
similarity=23.16605]
29: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.92567044,
final_local_scores=0.90464896, pem=0.0, similarity=31.08624]
30: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.88785225,
final_local_scores=0.8633383, pem=0.0, similarity=29.85345]
31: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.9417017,
final_local_scores=0.9221607, pem=0.0, similarity=31.608828]
32: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.96793646,
final_local_scores=0.95081824, pem=0.0, similarity=32.464027]
33: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.9707301,
final_local_scores=0.9538699, pem=0.0, similarity=32.55509]
34: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.9399889,
final_local_scores=0.9202897, pem=0.0, similarity=31.552992]
35: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.47538227,
final_local_scores=0.45039237, pem=-0.076093495,
similarity=17.944128]
36: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.44509703,
final_local_scores=0.4151615, pem=-0.1310191,
similarity=17.191462]
37: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.91834265,
final_local_scores=0.89664453, pem=0.0, similarity=30.847376]
38: 6886 Chicago
    [pem_pos=0, final_scores=0.6075531,
final_local_scores=0.646159, pem=-0.20658566,
similarity=24.495832]
39: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.6076651,
final_local_scores=0.5831686, pem=0.0, similarity=21.492628]
40: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.9474321,
final_local_scores=0.92842025, pem=0.0, similarity=31.795624]
41: 5407 California
    [pem_pos=0, final_scores=0.6378127,
```

```
final_local_scores=0.6895515, pem=-0.08662026,
similarity=25.13836]
42: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.4598388,
final_local_scores=0.43126458, pem=-0.1310191,
similarity=17.67201]
43: 343170 Pete_Wilson
    [pem_pos=0, final_scores=0.5071764,
final_local_scores=0.5088019, pem=0.0, similarity=19.273376]
44: 25433 Ronald_Reagan
    [pem_pos=0, final_scores=0.3735754,
final_local_scores=0.3257429, pem=0.0, similarity=13.810531]
45: 11955 George_H._W._Bush
    [pem_pos=1, final_scores=0.25814182,
final_local_scores=0.19916156, pem=-0.80923575,
similarity=14.433841]
46: 33057 White_House
    [pem_pos=0, final_scores=0.44919127,
final_local_scores=0.4433456, pem=-0.0056634876,
similarity=17.350828]
47: 343170 Pete_Wilson
    [pem_pos=0, final_scores=0.40668827,
final_local_scores=0.3990339, pem=0.0, similarity=15.99768]
```

Sample: b'735_BALANCE-Water'

BALANCE-Water Dist 1 Johnson Cty , Kan . , at \$ 11 mln .
WATER DISTRICT 1 OF JOHNSON CO . , KS RE : \$ 45,020,000 WATER
REVENUE BONDS \$ 22,040,000 SER . 1996A \$ 22,980,000 RFDG ,
SER . 1996B MOODY 'S : Aa S&P : AA+ Delivery Date :
09/05/1996 (FIRM) 06/01 12/01 MATURITY SER A SER B SER A
SER B

1998 665M 840M 570M 2000 - - 605M 2001 - 70M - 2002 - 895M
600M 2003 705M - 795M 2004 655M 90M 965M 2009 65M - - 2010
60M - 100M 2011 30M - 90M 2012 20M - 35M TOTAL : 11,450 A.G.
Edwards & Sons , Inc . -- U.S. Municipal Desk , 212-859-1650

Sample: b'736_CME'

[1 1 CME 1] lumber futures close lower on profit taking . [2
2 CHICAGO 2] 1996-08-28 Profit taking continued to weigh on [3
3 CME 3] lumber futures but prices ended only slightly lower
as strong cash markets underpinned futures , traders said .
The same pattern of the past few days persisted with futures
declining early on the profit taking before firming late .
There was cash-related buying late from people who want to
take delivery of the September contract , they said . Cash
sources noted that although the cash market is generally
quiet , prices remain firm on demand for prompt delivery
wood , they added . Random Lengths quoted cash spruce at \$
419 per tbf , up \$ 5 from last Friday and \$ 7 over the last
midweek quote . Reduced concern over Hurricane [4 4 Edouard 4]

prompted some of the early profit taking . Expectations the storm would turn more to the north partly eased concerns , they said . Lumber closed \$ 2.20 to \$ 0.20 per tbf lower with September off \$ 0.70 at \$ 413.20 and November off most at \$ 369.00 per tbf . -- Jerry Bieszk 312-408-8725

1: 351631 Chicago_Mercantile_Exchange
2: 6886 Chicago
3: 351631 Chicago_Mercantile_Exchange
4: 3589067 Hurricane_Edouard_(1996)
1: 351631 Chicago_Mercantile_Exchange
[pem_pos=1, final_scores=0.62071335,
final_local_scores=0.639266, pem=-1.04227,
similarity=28.834711]
2: 6886 Chicago
[pem_pos=0, final_scores=0.5646234,
final_local_scores=0.6635105, pem=-0.20658566,
similarity=25.013634]
3: 351631 Chicago_Mercantile_Exchange
[pem_pos=1, final_scores=0.55785805,
final_local_scores=0.5706061, pem=-1.04227,
similarity=26.785763]
4: 3589067 Hurricane_Edouard_(1996)
[pem_pos=0, final_scores=0.43262988,
final_local_scores=0.61269516, pem=-1.0413741,
similarity=28.036913]

Sample: b'737_WHEAT--Rains'
WHEAT--Rains boost [1 1 U.S. 1] HRW planting prospects . Greg Frost [2 2 KANSAS CITY 2] , Mo . 1996-08-28 Above-normal summer rainfall in the [3 3 U.S. 3] [4 4 High Plains 4] has produced near-ideal conditions for planting the 1997 hard red winter wheat crop , analysts said Wednesday . From central [5 5 Texas 5] north to [6 6 Kansas 6] , rains throughout July and August have relieved most of the drought conditions that plagued the region earlier this year . " Our moisture situation is excellent , especially for fall planting of winter wheat , " said Kim Anderson , extension wheat marketing economist at [7 7 Oklahoma State University 7] . The irony of the above-average summer rainfall was not lost on [8 8 High Plains 8] wheat producers , who only three months ago were caught in a drought so severe that old-timers likened conditions to the " [9 9 Dust Bowl 9] " days of the 1930s . " It 's definitely a turnaround from this past year , but you know last year we had pretty good moisture about this time of year , and then about October 1 it quit , " said Mark Hodges , executive director of the Oklahoma Wheat Commission . " Hopefully that 's not going to happen this year . " According to figures released by the Oklahoma Climatological Survey , an average of 20.19 inches fell across the state between March 1 and August 26 , 1996 . That 's about 1/2 inch above the average for the same time period , according to Howard Johnson , associate state

climatologist at the [10 10 University of Oklahoma 10] . He noted that the majority of that 20.19 inches had fallen since July . As an example of just how dry it was , data showed that between October 1 , 1995 and March 1 , 1996 , the state received an average of only 4.6 inches of rainfall . In northern [11 11 Texas 11] , the current rainfall situation was similar to most of [12 12 Oklahoma 12] , said Rodney Mosier , executive assistant for the Texas Wheat Producers . " Up here in the [13 13 Texas Panhandle 13] , we 've had some extremely beneficial rains that came through within the last several days and are really setting us up for ideal conditions for planting wheat , " Mosier said . But he warned that the situation was not as ideal in central and southern [1 4 14 Texas 14] , where moisture levels were still short despite the rains brought by Hurricane Dolly last week . In [1 5 15 Kansas 15] , typically the number one [16 16 U.S. 16] hard red winter wheat producer , topsoil moisture levels were rated mostly adequate during the week ended Sunday , according to the state 's agricultural statistics service . In its weekly report released Monday , the service said [17 17 Kansas 17] topsoil moisture was rated eight percent surplus , 77 percent adequate and 15 percent short to very short . [18 1 8 Oklahoma 18] 's Agricultural Statistics Service showed similar conditions , rating topsoil moisture levels as seven percent surplus , 81 percent adequate and 12 percent short to very short . Data on topsoil moisture ratings were not released by the Texas Agricultural Statistics Service . --

Greg Frost , 816 561-8671

1: 3434750 United_States
2: 17454 Kansas_City,_Missouri
3: 3434750 United_States
4: 2356523 High_Plains_(United_States)
5: 29810 Texas
6: 16716 Kansas
7: 323078 Oklahoma_State_University-Stillwater
8: 2356523 High_Plains_(United_States)
9: 59749 Dust_Bowl
10: 323072 University_of_Oklahoma
11: 29810 Texas
12: 22489 Oklahoma
13: 317673 Texas_Panhandle
14: 29810 Texas
15: 16716 Kansas
16: 3434750 United_States
17: 16716 Kansas
18: 22489 Oklahoma
1: 3434750 United_States
[pem_pos=0, final_scores=0.53335065,
final_local_scores=0.51553637, pem=-0.14075686,
similarity=20.239803]
2: 17454 Kansas_City,_Missouri
[pem_pos=0, final_scores=0.33027905,

```
final_local_scores=0.302193, pem=-0.8046767,
similarity=17.483711]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.5305889,
final_local_scores=0.51251966, pem=-0.14075686,
similarity=20.14978]
4: 2356523 High_Plains_(United_States)
    [pem_pos=0, final_scores=0.38244712,
final_local_scores=0.41879904, pem=0.0, similarity=16.587511]
5: 29810 Texas
    [pem_pos=0, final_scores=0.6999797,
final_local_scores=0.67208976, pem=-0.16191302,
similarity=25.02672]
6: 16716 Kansas
    [pem_pos=0, final_scores=0.5297522,
final_local_scores=0.4799825, pem=-0.3846871,
similarity=20.505337]
7: 323078 Oklahoma_State_University-Stillwater
    [pem_pos=0, final_scores=0.39855903,
final_local_scores=0.4198665, pem=0.0, similarity=16.619368]
8: 2356523 High_Plains_(United_States)
    [pem_pos=0, final_scores=0.4321136,
final_local_scores=0.46637338, pem=0.0, similarity=18.007225]
9: 59749 Dust_Bowl
    [pem_pos=0, final_scores=0.48247808,
final_local_scores=0.5329241, pem=0.0, similarity=19.99323]
10: 323072 University_of_Oklahoma
    [pem_pos=0, final_scores=0.51064354,
final_local_scores=0.5342717, pem=0.0, similarity=20.033447]
11: 29810 Texas
    [pem_pos=0, final_scores=0.7085921,
final_local_scores=0.68149745, pem=-0.16191302,
similarity=25.307465]
12: 22489 Oklahoma
    [pem_pos=0, final_scores=0.58568805,
final_local_scores=0.535651, pem=-0.35266355,
similarity=21.992447]
13: 317673 Texas_Panhandle
    [pem_pos=0, final_scores=0.47258574,
final_local_scores=0.4605919, pem=0.0, similarity=17.834694]
14: 29810 Texas
    [pem_pos=0, final_scores=0.73265606,
final_local_scores=0.7077837, pem=-0.16191302,
similarity=26.091898]
15: 16716 Kansas
    [pem_pos=0, final_scores=0.44647777,
final_local_scores=0.38901782, pem=-0.3846871,
similarity=17.79077]
16: 3434750 United_States
    [pem_pos=0, final_scores=0.52601093,
final_local_scores=0.5075188, pem=-0.14075686,
similarity=20.000546]
```

```
17: 16716 Kansas
    [pem_pos=0, final_scores=0.5302097,
final_local_scores=0.48048228, pem=-0.3846871,
similarity=20.52025]
18: 22489 Oklahoma
    [pem_pos=0, final_scores=0.5684358,
final_local_scores=0.5168056, pem=-0.35266355,
similarity=21.430061]
```

Sample: b'738_First'

First Union National Bank of Fla. settles suit . [1 1 JACKSONVILLE 1] , Fla. 1996-08-28 First Union National Bank of Florida said on Wednesday it agreed to settle a class action law suit involving its collateral protection insurance (CPI) program . To provide for the settlements , First Union has established a common fund of \$ 4.7 million for cash refunds and \$ 19.4 million in credit refunds for outstanding CPI balances . The bank is a division of First Union Corp . The bank said most of the charges resulted from loan portfolios from banks and thrifts that were acquired in the 1980s . First Union said it has discontinued CPI as an element of its motor vehicle or boat installment loan contracts . As part of the settlement agreement , customers who had CPI placed on loans from January 1 , 1986 to September 31 , 1996 , will receive cash or credit refunds , the bank said . Cash refunds will go to those who paid their loans to First Union while credit refunds will go to those who have existing loan balances , the bank said .

```
1: 60613 Jacksonville,_Florida
```

```
1: 60613 Jacksonville,_Florida
```

```
    [pem_pos=0, final_scores=0.15664347,
final_local_scores=0.44369638, pem=-0.5613812,
similarity=20.383377]
```

Sample: b'739_Amoco'

[1 1 Amoco 1] says in talks over [2 2 Yemen 2] oil acreage . [3 3 NEW YORK 3] 1996-08-28 Amoco Corp officials said the company is in talks over crude oil production sharing in [4 4 Yemen 4] , but declined to comment on a published report [5 5 Amoco 5] had reached preliminary agreement on a block in the Shabwa area . " We 've been asked (by [6 6 Yemen 6]) not to comment (on the talks) , " said [7 7 Amoco 7] spokesman Dan Dietsch . " We can neither confirm nor deny that report , " he said . According to [8 8 Middle East Economic Survey 8] (MEES) , [9 9 Yemen 9] and [10 10 Amoco 10] signed a " memorandum of understanding " for a production-sharing agreement in Shabwa Block No . S-1 in the former [11 11 South Yemen 11] , which united with [12 12 North Yemen 12] in 1990 . According to another [13 13 Amoco 13] official , the company is not exploring now for oil anywhere in [14 14 Yemen 14] . The officials said [15 15 Amoco 15] was deferring to the [16 16 Yemeni 16] Ministry of Petroleum and Minerals for

any specific comments on the Amoco-Yemen talks . The former [1
7 17 Soviet Union 17] was displaced as contractor of the
potentially rich Shabwa oilfields once it collapsed in
December 1991 , according to the International Petroleum
Encyclopedia . -- Oliver Ludwig , New York Energy Desk +1 212
859 1633

1: 59418 Amoco
2: 350939 Yemen
3: 645042 New_York_City
4: 350939 Yemen
5: 59418 Amoco
6: 350939 Yemen
7: 59418 Amoco
8: 27308038 Middle_East_Economic_Survey
9: 350939 Yemen
10: 59418 Amoco
11: 718749 South_Yemen
12: 720778 North_Yemen
13: 59418 Amoco
14: 350939 Yemen
15: 59418 Amoco
16: 350939 Yemen
17: 26779 Soviet_Union
1: 59418 Amoco
[pem_pos=0, final_scores=0.61730415,
final_local_scores=0.6466719, pem=0.0, similarity=23.387691]
2: 350939 Yemen
[pem_pos=0, final_scores=0.61497074,
final_local_scores=0.610596, pem=-0.27215755,
similarity=23.791151]
3: 645042 New_York_City
[pem_pos=1, final_scores=0.36455166,
final_local_scores=0.41568148, pem=-1.7906687,
similarity=26.232412]
4: 350939 Yemen
[pem_pos=0, final_scores=0.49779058,
final_local_scores=0.48259443, pem=-0.27215755,
similarity=19.971329]
5: 59418 Amoco
[pem_pos=0, final_scores=0.58185095,
final_local_scores=0.60794467, pem=0.0, similarity=22.231995]
6: 350939 Yemen
[pem_pos=0, final_scores=0.5990279,
final_local_scores=0.5931808, pem=-0.27215755,
similarity=23.271446]
7: 59418 Amoco
[pem_pos=0, final_scores=0.5934451,
final_local_scores=0.6206095, pem=0.0, similarity=22.60994]
8: 27308038 Middle_East_Economic_Survey
[pem_pos=0, final_scores=0.34723932,
final_local_scores=0.4943384, pem=0.0, similarity=18.841757]
9: 350939 Yemen

```
[pem_pos=0, final_scores=0.5728376,
final_local_scores=0.5645719, pem=-0.27215755,
similarity=22.4177]
10: 59418 Amoco
    [pem_pos=0, final_scores=0.5625902,
final_local_scores=0.5869053, pem=0.0, similarity=21.604137]
11: 718749 South_Yemen
    [pem_pos=0, final_scores=0.37792104,
final_local_scores=0.38507813, pem=0.0, similarity=15.581213]
12: 720778 North_Yemen
    [pem_pos=0, final_scores=0.25500613,
final_local_scores=0.2892703, pem=-0.028031487,
similarity=12.874556]
13: 59418 Amoco
    [pem_pos=0, final_scores=0.6588115,
final_local_scores=0.6920125, pem=0.0, similarity=24.740746]
14: 350939 Yemen
    [pem_pos=0, final_scores=0.50197786,
final_local_scores=0.48716843, pem=-0.27215755,
similarity=20.107826]
15: 59418 Amoco
    [pem_pos=0, final_scores=0.56642646,
final_local_scores=0.59109575, pem=0.0, similarity=21.729189]
16: 350939 Yemen
    [pem_pos=0, final_scores=0.59495866,
final_local_scores=0.5887358, pem=-0.006034785,
similarity=21.691582]
17: 26779 Soviet_Union
    [pem_pos=0, final_scores=0.36373156,
final_local_scores=0.3938272, pem=-0.075323656,
similarity=16.251925]
```

Sample: b'740_Burundi'

[1 1 Burundi 1] defends military regime to hostile UN . Evelyn Leopold [2 2 UNITED NATIONS 2] 1996-08-28 [3 3 Burundi 3] 's ambassador on Wednesday lashed out at economic sanctions imposed by [4 4 African 4] states and said any thought of an arms embargo would be a windfall for guerrillas fighting his army-run government . In a lengthy debate on [5 5 Burundi 5] before the [6 6 U.N. Security Council 6] , Ambassador Nsanze Terence said the new military government took over to stabilise the country and wanted negotiations under former [7 7 Tanzanian 7] President [8 8 Julius Nyrere 8] . Nearly every [9 9 African 9] member who spoke , as well as most [10 10 Security Council 10] members , however , were unsympathetic towards the government of President [11 11 Pierre Buyoya 11] , an army major put in power in a July coup by the Tutsi-run military , which is locked in a guerrilla war with the majority Hutus . " These ([12 12 African 12]) brothers should have been the first to bind the wounds of [13 13 Burundi 13] , " Terence said of the economic embargo . " Quite the contrary , [14 14 Burundi 14] has seen economic war

declared against it by fellow [15 15 African 15] people ... a gratuitous immolation of the people of [16 16 Burundi 16] . " He said his government had just asked [17 17 U.N. 17] human rights monitors to increase their numbers in [18 18 Burundi 18] in an effort " to put an end to this vicious circle of violence . " More than 150,000 people have been killed in violence between the minority [19 19 Tutsis 19] and the majority Hutus since 1993 . [20 20 Botswana 20] 's envoy , Mothusi Nkgowe , said coups should be relegated " to the dump heap of history " as there could be no justification for the overthrow of a legitimate government . [21 21 Chile 21] has proposed a resolution , still under discussion , that would impose an immediate arms embargo on [22 22 Burundi 22] and call for negotiations . The draft suggests further sanctions against those who impede a political solution . Among the council 's five permanent members , [23 23 Russia 23] and the [24 24 United States 24] appeared to support most elements of the [25 25 Chilean 25] proposal , while [26 26 Britain 26] , [27 27 France 27] and [28 28 China 28] were cautious .

Terence , a [29 29 Tutsi 29] , said any arms embargo would leave the army unable to defend itself against [30 30 Hutu 30] guerrillas and leave the population exposed to " armed terroristts . " But [31 31 Chilean 31] Ambassador [32 32 Juan Somavia 32] said : " Every weapon that reached [33 33 Burundi 33] is a weapon aimed mainly at killing an unarmed civilian . We must not send a signal different from the [34 34 African 34] leaders themselves . Inaction is becoming the worst possible course of action . " [35 35 Burundi 35] 's parliament has been suspended and political parties are banned but Terence told reporters [36 36 Buyoya 36] would reconvene a new type of national assembly in October . The [37 37 United States 37] said the coup leaders had taken no steps to restore democracy and indiscriminate killings continued . Ambassador [38 38 Karl Inderfurth 38] said the new government should have " unconditional " negotiations with all parties inside and outside of the country . He said [39 39 Washington 39] strongly supported the economic sanctions imposed already and if these did not work the council would consider " an arms embargo or targeted sanctions against faction leaders . " But he said the international community had to be prepared for the worst and avoid a replay of the horrors in neighbouring [40 40 Rwanda 40] , where widespread genocide broke out against the [41 41 Tutsis 41] two years ago . He again said the [42 42 United Nations 42] should draw up contingency plans for a rapid humanitarian intervention .

1: 21490998 Burundi

2: 31769 United_Nations

3: 21490998 Burundi

4: 5334607 Africa

5: 21490998 Burundi

6: 31956 United_Nations_Security_Council

7: 30118 Tanzania

```
8: 358637 Julius_Nyerere
9: 5334607 Africa
10: 31956 United_Nations_Security_Council
11: 616447 Pierre_Buyoya
12: 5334607 Africa
13: 21490998 Burundi
14: 21490998 Burundi
15: 5334607 Africa
16: 21490998 Burundi
17: 31769 United_Nations
18: 21490998 Burundi
19: 67050 Tutsi
20: 3464 Botswana
21: 5489 Chile
22: 21490998 Burundi
23: 25391 Russia
24: 3434750 United_States
25: 5489 Chile
26: 31717 United_Kingdom
27: 5843419 France
28: 5405 China
29: 67050 Tutsi
30: 59537 Hutu
31: 5489 Chile
32: 5268497 Juan_Somavía
33: 21490998 Burundi
34: 5334607 Africa
35: 21490998 Burundi
36: 616447 Pierre_Buyoya
37: 3434750 United_States
38: 8334599 Karl_Inderfurth
39: 108956 Washington,_D.C.
40: 25645 Rwanda
41: 67050 Tutsi
42: 31769 United_Nations
1: 21490998 Burundi
    [pem_pos=0, final_scores=0.4744494,
final_local_scores=0.4262305, pem=-0.3738042,
similarity=18.842085]
2: 31769 United_Nations
    [pem_pos=0, final_scores=0.43072045,
final_local_scores=0.419142, pem=0.0, similarity=16.597746]
3: 21490998 Burundi
    [pem_pos=0, final_scores=0.45808464,
final_local_scores=0.40835446, pem=-0.3738042,
similarity=18.30863]
4: 5334607 Africa
    [pem_pos=0, final_scores=0.5960501,
final_local_scores=0.57963264, pem=-0.91165566,
similarity=26.344833]
5: 21490998 Burundi
    [pem_pos=0, final_scores=0.54782706,
```

```
final_local_scores=0.5063845, pem=-0.3738042,
similarity=21.234041]
6: 31956 United_Nations_Security_Council
    [pem_pos=0, final_scores=0.36483574,
final_local_scores=0.39140368, pem=0.0, similarity=15.769979]
7: 30118 Tanzania
    [pem_pos=0, final_scores=0.50984234,
final_local_scores=0.4965014, pem=0.0, similarity=18.906305]
8: 358637 Julius_Nyerere
    [pem_pos=0, final_scores=0.75835544,
final_local_scores=0.82645863, pem=0.0, similarity=28.752888]
9: 5334607 Africa
    [pem_pos=0, final_scores=0.6356116,
final_local_scores=0.6228476, pem=-0.91165566,
similarity=27.634453]
10: 31956 United_Nations_Security_Council
    [pem_pos=0, final_scores=0.37815237,
final_local_scores=0.40595007, pem=0.0, similarity=16.204073]
11: 616447 Pierre_Buyoya
    [pem_pos=0, final_scores=0.5724305,
final_local_scores=0.67214817, pem=0.0, similarity=24.147955]
12: 5334607 Africa
    [pem_pos=0, final_scores=0.5556046,
final_local_scores=0.53545207, pem=-0.91165566,
similarity=25.026396]
13: 21490998 Burundi
    [pem_pos=0, final_scores=0.50521266,
final_local_scores=0.4598347, pem=-0.3738042,
similarity=19.844902]
14: 21490998 Burundi
    [pem_pos=0, final_scores=0.494085,
final_local_scores=0.44767946, pem=-0.3738042,
similarity=19.482166]
15: 5334607 Africa
    [pem_pos=0, final_scores=0.5608392,
final_local_scores=0.54117006, pem=-0.91165566,
similarity=25.197033]
16: 21490998 Burundi
    [pem_pos=0, final_scores=0.5060046,
final_local_scores=0.46069974, pem=-0.3738042,
similarity=19.870716]
17: 31769 United_Nations
    [pem_pos=0, final_scores=0.5448254,
final_local_scores=0.5437843, pem=0.0, similarity=20.317322]
18: 21490998 Burundi
    [pem_pos=0, final_scores=0.5030806,
final_local_scores=0.45750576, pem=-0.3738042,
similarity=19.775402]
19: 67050 Tutsi
    [pem_pos=0, final_scores=0.41889662,
final_local_scores=0.44120586, pem=0.0, similarity=17.256176]
20: 3464 Botswana
```

```
[pem_pos=0, final_scores=0.55248076,
final_local_scores=0.5375053, pem=-0.37041438,
similarity=22.144314]
21: 5489 Chile
    [pem_pos=0, final_scores=0.6751419,
final_local_scores=0.6979584, pem=-0.27404457,
similarity=26.40848]
22: 21490998 Burundi
    [pem_pos=0, final_scores=0.52351326,
final_local_scores=0.47982532, pem=-0.3738042,
similarity=20.441462]
23: 25391 Russia
    [pem_pos=0, final_scores=0.6376199,
final_local_scores=0.6776241, pem=-0.20299491,
similarity=25.415285]
24: 3434750 United_States
    [pem_pos=0, final_scores=0.48915774,
final_local_scores=0.49223548, pem=-0.073457584,
similarity=19.178476]
25: 5489 Chile
    [pem_pos=0, final_scores=0.54351693,
final_local_scores=0.5541782, pem=-0.25753683,
similarity=22.028019]
26: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.4931928,
final_local_scores=0.49895138, pem=-0.61114657,
similarity=22.302929]
27: 5843419 France
    [pem_pos=0, final_scores=0.5809317,
final_local_scores=0.5880606, pem=-0.16240834,
similarity=22.521816]
28: 5405 China
    [pem_pos=0, final_scores=0.510556,
final_local_scores=0.5129854, pem=-0.15669827,
similarity=20.25037]
29: 67050 Tutsi
    [pem_pos=0, final_scores=0.4892183,
final_local_scores=0.5180216, pem=0.0, similarity=19.54851]
30: 59537 Hutu
    [pem_pos=0, final_scores=0.5042553,
final_local_scores=0.54436266, pem=0.0, similarity=20.334581]
31: 5489 Chile
    [pem_pos=0, final_scores=0.5783443,
final_local_scores=0.5922218, pem=-0.25753683,
similarity=23.163319]
32: 5268497 Juan_Somavía
    [pem_pos=0, final_scores=0.49987954,
final_local_scores=0.6060381, pem=0.0, similarity=22.175098]
33: 21490998 Burundi
    [pem_pos=0, final_scores=0.5307188,
final_local_scores=0.4876963, pem=-0.3738042,
similarity=20.676348]
```

```
34: 5334607 Africa
    [pem_pos=0, final_scores=0.5217258,
final_local_scores=0.49844462, pem=-0.91165566,
similarity=23.92202]
35: 21490998 Burundi
    [pem_pos=0, final_scores=0.5533913,
final_local_scores=0.51246256, pem=-0.3738042,
similarity=21.415422]
36: 616447 Pierre_Buyoya
    [pem_pos=0, final_scores=0.5125627,
final_local_scores=0.6067517, pem=0.0, similarity=22.196394]
37: 3434750 United_States
    [pem_pos=0, final_scores=0.46559823,
final_local_scores=0.46650034, pem=-0.073457584,
similarity=18.410488]
38: 8334599 Karl_Inderfurth
    [pem_pos=0, final_scores=0.5298278,
final_local_scores=0.6657908, pem=0.0, similarity=23.958239]
39: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.4070869,
final_local_scores=0.4343714, pem=-1.7496481,
similarity=26.56708]
40: 25645 Rwanda
    [pem_pos=0, final_scores=0.49952227,
final_local_scores=0.45109946, pem=-0.42949,
similarity=19.887054]
41: 67050 Tutsi
    [pem_pos=0, final_scores=0.4799227,
final_local_scores=0.5078676, pem=0.0, similarity=19.245495]
42: 31769 United_Nations
    [pem_pos=0, final_scores=0.46723467,
final_local_scores=0.4590283, pem=0.0, similarity=17.788033]
```

Sample: b'741_Swiss'

```
[1 1 Swiss 1] arrest [2 2 Rwandan 2] on genocide suspicion .
[3 3 BERNE 3] 1996-08-28 [4 4 Swiss 4] authorities said on
Wednesday they had arrested a former [5 5 Rwandan 5] mayor ,
now living in [6 6 Switzerland 6] , on suspicion of violating
human rights during the genocide in his country in 1994 . The
Defence Ministry said in a statement that investigations were
still in the preliminary stage but it was cooperating closely
with police in the cantons of [7 7 Geneva 7] and [1 1 Freiburg
1] . It did not identify the man .
1: 26748 Switzerland
2: 25645 Rwanda
3: 49749 Bern
4: 26748 Switzerland
5: 25645 Rwanda
6: 26748 Switzerland
7: 12521 Geneva
1: 164065 Fribourg
    [winner: 101832 Freiburg_im_Breisgau, pem_pos=0,
```

```

final_scores=0.13758971, final_local_scores=0.10923223,
pem=-0.20749207, similarity=8.477795]
    gt: 164065 Fribourg, pem_pos=6, final_scores=0.10842012,
final_local_scores=0.0795871, pem=-2.2581654,
similarity=18.745005 ]
1: 26748 Switzerland
    [pem_pos=1, final_scores=0.831513,
final_local_scores=0.80346024, pem=-0.1392413,
similarity=28.823786]
2: 25645 Rwanda
    [pem_pos=0, final_scores=0.5602481,
final_local_scores=0.5868431, pem=0.0, similarity=21.602282]
3: 49749 Bern
    [pem_pos=0, final_scores=0.41163725,
final_local_scores=0.36421382, pem=-0.3095387,
similarity=16.6419]
4: 26748 Switzerland
    [pem_pos=0, final_scores=0.8897057,
final_local_scores=0.867027, pem=-0.1392413,
similarity=30.720743]
5: 25645 Rwanda
    [pem_pos=0, final_scores=0.44027144,
final_local_scores=0.45578682, pem=0.0, similarity=17.691301]
6: 26748 Switzerland
    [pem_pos=0, final_scores=0.5867702,
final_local_scores=0.5361159, pem=-0.18598008,
similarity=21.099869]
7: 12521 Geneva
    [pem_pos=0, final_scores=0.51936847,
final_local_scores=0.4946068, pem=-0.24772356,
similarity=20.196926]
1: 101832 Freiburg_im_Breisgau
    [pem_pos=0, final_scores=0.13758971,
final_local_scores=0.10923223, pem=-0.20749207,
similarity=8.477795]

```

Sample: b'742_OPTIONS'

```

OPTIONS -- EOE options volumes - close . [1 1 AMSTERDAM 1]
1996-08-28 1605 [2 2 GMT 2] CALLS PUTS PCT OF TOTAL TOTAL
VOLUME -- 83,008 60,131 22,877 -- FEATURES - [3 3 AEX 3]
INDEX 7,391 5,658 15.72 - [4 4 AHOLD 4] 7,190 1,123 10.01 -
BOLSWESSANEN 4,420 705 6.17 - [5 5 ABN AMRO 5] 3,003 1,940
5.95 - [6 6 ING 6] 3,853 673 5.45 - VNU 3,060 843 4.70 -- [7 7
Amsterdam 7] newsdesk +31 20 504 5000 ( Fax 020-504-5040 )
1: 844 Amsterdam
2: 12701 Greenwich_Mean_Time
3: 1371940 AEX_index
4: 735427 Ahold
5: 383976 ABN_AMRO
6: 230965 ING_Group
7: 844 Amsterdam
1: 844 Amsterdam

```

```
[pem_pos=0, final_scores=0.63529515,
final_local_scores=0.65990007, pem=-0.0642803,
similarity=24.132013]
2: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.6151233,
final_local_scores=0.75251, pem=0.0, similarity=26.546112]
3: 1371940 AEX_index
    [pem_pos=0, final_scores=0.38249266,
final_local_scores=0.47065276, pem=0.0, similarity=18.13493]
4: 735427 Ahold
    [pem_pos=0, final_scores=0.45010328,
final_local_scores=0.5124877, pem=0.0, similarity=19.383368]
5: 383976 ABN_AMRO
    [pem_pos=0, final_scores=0.49872077,
final_local_scores=0.52301437, pem=0.0, similarity=19.697504]
6: 230965 ING_Group
    [pem_pos=0, final_scores=0.50690645,
final_local_scores=0.53380394, pem=0.0, similarity=20.019487]
7: 844 Amsterdam
    [pem_pos=0, final_scores=0.596823,
final_local_scores=0.61787504, pem=-0.0642803,
similarity=22.877903]
```

Sample: b'743_French'

```
[1 1 French 1] tax office sucks in money . [2 2 PARIS 2]
1996-08-28 Workers fixing the ceiling of a tax office in [3 3
Paris 3] found a dozen seven-year-old cheques for a total of
six million francs ( $ 1.2 million ) in a ventilation pipe ,
the weekly [4 4 Le Canard Enchaine 4] said on Wednesday . A
Finance Ministry official explained that the cheques for
corporate tax payments had been sucked into the ventilation
system , the weekly reported . The companies had been
contacted at the time and had not been fined for failing to
pay , the official said .
```

```
1: 5843419 France
2: 22989 Paris
3: 22989 Paris
4: 297545 Le_Canard_enchaîné
1: 5843419 France
    [pem_pos=0, final_scores=0.7863248,
final_local_scores=0.7986093, pem=-0.80223274,
similarity=32.284473]
2: 22989 Paris
    [pem_pos=0, final_scores=0.606539,
final_local_scores=0.5664142, pem=-0.115301244,
similarity=21.619667]
3: 22989 Paris
    [pem_pos=0, final_scores=0.6096045,
final_local_scores=0.5697627, pem=-0.115301244,
similarity=21.719595]
4: 297545 Le_Canard_enchaîné
    [pem_pos=0, final_scores=0.4986391,
```

```
final_local_scores=0.5901773, pem=0.0, similarity=21.70178]
```

Sample: b'744_Swiss'

[1 1 Swiss 1] bond market closing report . [2 2 ZURICH 2]
1996-08-28 [3 3 Swiss 3] bonds ended mostly higher in
generally quiet activity , with the September confederate
bond futures contract holding just above 113.00 . " Today was
very quiet after a lot of activity on Tuesday , " said one [4
4 Swiss 4] bond futures trader . He said the market began
strong , gave up some gains at midday and then was able to
recover back at the close , but all on light volume . In the
primary market , Eksportfinans and Suedwest LB launched
issues for 100 million and 300 million [5 5 Swiss 5] francs ,
respectively . Eksportfinans ' seven-year issue was quoted at
a yield of 3.98 percent , and Suedwest LB 's five-year issue
was quoted at 3.60 percent . [6 6 Swiss 6] money market rates
remained lower at around 1.75 percent offered . As to
fundamentals , economists at [7 7 Credit Suisse 7] said they
expect the country 's gross domestic product to be flat to
negative in 1996 , and to grow only 0.6 percent in 1997 . The
[8 8 Swiss 8] government also said Wednesday it had made
progress in cutting [9 9 Switzerland 9] 's projected spending
for 1997 . [10 10 Switzerland 10] reports July consumer
prices later this week . Closing prices as follows : Sept
conf bond up 12 at 113.02 . Sept comi medium-term bond up
three at 109.45 . Sept Euro Swiss francs up three at 97.82 .
-- Cash : 4-1/2 Apr 2006 bond 101.80 / 90 yield 4.252 pct --
Zurich Editorial , +41 1 631 7340

1: 26748 Switzerland

2: 40334603 Zürich

3: 26748 Switzerland

4: 26748 Switzerland

5: 26748 Switzerland

6: 26748 Switzerland

7: 842695 Credit_Suisse

8: 26748 Switzerland

9: 26748 Switzerland

10: 26748 Switzerland

1: 26748 Switzerland

[pem_pos=0, final_scores=0.7122731,
final_local_scores=0.6531468, pem=-0.1392413,
similarity=24.338133]

2: 40334603 Zürich

[pem_pos=0, final_scores=0.6459604,
final_local_scores=0.62515014, pem=-0.08839822,
similarity=23.226162]

3: 26748 Switzerland

[pem_pos=0, final_scores=0.82812554,
final_local_scores=0.77969795, pem=-0.1392413,
similarity=28.114672]

4: 26748 Switzerland

[pem_pos=0, final_scores=0.8042293,

```
final_local_scores=0.753595, pem=-0.1392413,
similarity=27.335707]
5: 26748 Switzerland
    [pem_pos=0, final_scores=0.67576694,
final_local_scores=0.6132693, pem=-0.1392413,
similarity=23.14811]
6: 26748 Switzerland
    [pem_pos=0, final_scores=0.7179579,
final_local_scores=0.6593566, pem=-0.1392413,
similarity=24.523445]
7: 842695 Credit_Suisse
    [pem_pos=0, final_scores=0.40952623,
final_local_scores=0.5022621, pem=0.0, similarity=19.078217]
8: 26748 Switzerland
    [pem_pos=0, final_scores=0.8210122,
final_local_scores=0.7719277, pem=-0.1392413,
similarity=27.882792]
9: 26748 Switzerland
    [pem_pos=0, final_scores=0.6240849,
final_local_scores=0.55681455, pem=-0.18598008,
similarity=21.717558]
10: 26748 Switzerland
    [pem_pos=0, final_scores=0.5994006,
final_local_scores=0.52985066, pem=-0.18598008,
similarity=20.9129]
```

Sample: b'745_Anti-Bhutto'

Anti-Bhutto rally draws about 8,000 in [1 1 Karachi 1] . [2 2 KARACHI 2] , [3 3 Pakistan 3] 1996-08-28 About 8,000 protesters marched through [4 4 Karachi 4] on Wednesday demanding the removal of [5 5 Pakistani 5] Prime Minister [6 6 Benazir Bhutto 6] , witnesses said . " From here we will march to [7 7 Islamabad 7] and by [8 8 God 8] we will not let [9 9 Benazir 9] and ([10 10 Bhutto 10] 's husband Asif Ali) [11 11 Zardari 11] escape justice , " [12 12 Nawaz Sharif 12] , leader of the main opposition [13 13 Pakistan Muslim League 13] told a rally organised by a 16-party alliance . [14 14 Sharif 14] accused [15 15 Bhutto 15] of corruption and nepotism , charges she has denied in the past . Witnesses said protesters carrying colourful party flags walked for several miles , chanting anti-government slogans . The event was part of an opposition campaign launched on August 14 , [16 16 Pakistan 16] 's independence day . [17 17 Sharif 17] said similar rallies would be held in the [18 18 Balochistan 18] provincial capital [19 19 Quetta 19] and the [20 20 Punjab 20] provincial capital [21 21 Lahore 21] before an opposition march on the capital [22 22 Islamabad 22] . " I promise the people of [23 23 Karachi 23] that those responsible for the extra-judicial killing of innocent youths would not be spared , " [24 24 Sharif 24] said . [25 25 Karachi 25] 's ethnic Mohajir National Movement ([26 26 MQM 26]) accuses the government of killing many of its militants in cold

blood . The government has denied the charge and blames the [2
7 27 MQM 27] for much of the violence that killed 2,000 people in the city last year . Political observers said the turn-out was disappointing for a city of about 12 million people , possibly indicating that the [28 28 MQM 28] , although a member of the opposition alliance , had not mobilised its supporters for the event . The turbulent southern port has been calmer this year , but police say more than 300 people have died in political unrest . The [29 29 MQM 29] speaks for Urdu-speaking Moslems who migrated from [30 30 India 30] at [31 31 Partition 31] in 1947 and their descendants . [32 32 Sharif 32] , a former prime minister , is the main political rival of [33 33 Bhutto 33] , who defeated him in the October 1993 election . He said only the removal of the government and an early election could save [34 34 Pakistan 34] from disaster . " We will dislodge the [35 35 Bhutto 35] government . It is a holy war for us , " he said . [36 36 Bhutto 36] has vowed to complete her five-year term .

- 1: 17123 Karachi
- 2: 17123 Karachi
- 3: 23235 Pakistan
- 4: 17123 Karachi
- 5: 23235 Pakistan
- 6: 193737 Benazir_Bhutto
- 7: 51153 Islamabad
- 8: 5042765 God
- 9: 193737 Benazir_Bhutto
- 10: 193737 Benazir_Bhutto
- 11: 993655 Asif_Ali_Zardari
- 12: 374012 Nawaz_Shari
- 13: 409053 Pakistan_Muslim_League_(N)
- 14: 374012 Nawaz_Shari
- 15: 193737 Benazir_Bhutto
- 16: 23235 Pakistan
- 17: 374012 Nawaz_Shari
- 18: 4687 Balochistan,_Pakistan
- 19: 93764 Quetta
- 20: 24751 Punjab,_Pakistan
- 21: 125315 Lahore
- 22: 51153 Islamabad
- 23: 17123 Karachi
- 24: 374012 Nawaz_Shari
- 25: 17123 Karachi
- 26: 1030551 Muttahida_Qaumi_Movement
- 27: 1030551 Muttahida_Qaumi_Movement
- 28: 1030551 Muttahida_Qaumi_Movement
- 29: 1030551 Muttahida_Qaumi_Movement
- 30: 14533 India
- 31: 265059 Partition_of_India
- 32: 374012 Nawaz_Shari
- 33: 193737 Benazir_Bhutto
- 34: 23235 Pakistan

35: 193737 Benazir_Bhutto
36: 193737 Benazir_Bhutto
1: 17123 Karachi
 [pem_pos=0, final_scores=0.40798432,
final_local_scores=0.38016325, pem=0.0, similarity=15.434543]
2: 17123 Karachi
 [pem_pos=0, final_scores=0.4896664,
final_local_scores=0.46938854, pem=0.0, similarity=18.097204]
3: 23235 Pakistan
 [pem_pos=0, final_scores=0.64355785,
final_local_scores=0.6364013, pem=-0.21944217,
similarity=24.274557]
4: 17123 Karachi
 [pem_pos=0, final_scores=0.3961479,
final_local_scores=0.36723375, pem=0.0, similarity=15.0487]
5: 23235 Pakistan
 [pem_pos=0, final_scores=0.5990389,
final_local_scores=0.5877711, pem=-0.49826378,
similarity=24.339611]
6: 193737 Benazir_Bhutto
 [pem_pos=0, final_scores=0.54525554,
final_local_scores=0.50038606, pem=0.0, similarity=19.02223]
7: 51153 Islamabad
 [pem_pos=0, final_scores=0.36796772,
final_local_scores=0.33367404, pem=0.0, similarity=14.047212]
8: 5042765 God
 [pem_pos=0, final_scores=0.526221,
final_local_scores=0.67209643, pem=-0.18657415,
similarity=25.161032]
9: 193737 Benazir_Bhutto
 [pem_pos=0, final_scores=0.4872825,
final_local_scores=0.43705934, pem=0.0, similarity=17.132435]
10: 193737 Benazir_Bhutto
 [pem_pos=0, final_scores=0.5445675,
final_local_scores=0.50128484, pem=0.0, similarity=19.049053]
11: 993655 Asif_Ali_Zardari
 [pem_pos=0, final_scores=0.43110824,
final_local_scores=0.40251172, pem=0.0, similarity=16.101465]
12: 374012 Nawaz_Shariif
 [pem_pos=0, final_scores=0.61250097,
final_local_scores=0.5781874, pem=0.0, similarity=21.343979]
13: 409053 Pakistan_Muslim_League_(N)
 [pem_pos=1, final_scores=0.18142228,
final_local_scores=0.14027965, pem=-1.3418244,
similarity=15.5729885]
14: 374012 Nawaz_Shariif
 [pem_pos=0, final_scores=1.0067997,
final_local_scores=1.0107558, pem=0.0, similarity=34.25268]
15: 193737 Benazir_Bhutto
 [pem_pos=0, final_scores=0.5954656,
final_local_scores=0.5568833, pem=0.0, similarity=20.708221]
16: 23235 Pakistan

```
[pem_pos=0, final_scores=0.5638884,
final_local_scores=0.54937446, pem=-0.21944217,
similarity=21.677504]
17: 374012 Nawaz_Sharif
    [pem_pos=0, final_scores=0.9918137,
final_local_scores=0.99438566, pem=0.0, similarity=33.764164]
18: 4687 Balochistan,_Pakistan
    [pem_pos=0, final_scores=0.26454204,
final_local_scores=0.25029033, pem=-0.1729325,
similarity=12.49931]
19: 93764 Quetta
    [pem_pos=0, final_scores=0.37882775,
final_local_scores=0.3551767, pem=0.0, similarity=14.688893]
20: 24751 Punjab,_Pakistan
    [pem_pos=1, final_scores=0.28051442,
final_local_scores=0.25897765, pem=-1.088274,
similarity=17.736326]
21: 125315 Lahore
    [pem_pos=0, final_scores=0.43169945,
final_local_scores=0.4102714, pem=0.0, similarity=16.33303]
22: 51153 Islamabad
    [pem_pos=0, final_scores=0.3490799,
final_local_scores=0.31304196, pem=0.0, similarity=13.431509]
23: 17123 Karachi
    [pem_pos=0, final_scores=0.3737836,
final_local_scores=0.34280413, pem=0.0, similarity=14.319672]
24: 374012 Nawaz_Sharif
    [pem_pos=0, final_scores=0.9309109,
final_local_scores=0.92785853, pem=0.0, similarity=31.778864]
25: 17123 Karachi
    [pem_pos=0, final_scores=0.41007257,
final_local_scores=0.38244432, pem=0.0, similarity=15.502614]
26: 1030551 Muttahida_Qaumi_Movement
    [pem_pos=0, final_scores=0.5205185,
final_local_scores=0.5076278, pem=0.0, similarity=19.238338]
27: 1030551 Muttahida_Qaumi_Movement
    [pem_pos=0, final_scores=0.49104947,
final_local_scores=0.4754374, pem=0.0, similarity=18.277714]
28: 1030551 Muttahida_Qaumi_Movement
    [pem_pos=0, final_scores=0.4995814,
final_local_scores=0.48475724, pem=0.0, similarity=18.555836]
29: 1030551 Muttahida_Qaumi_Movement
    [pem_pos=0, final_scores=0.44660634,
final_local_scores=0.42688996, pem=0.0, similarity=16.82896]
30: 14533 India
    [pem_pos=0, final_scores=0.53074384,
final_local_scores=0.5889999, pem=-0.13622136,
similarity=22.407438]
31: 265059 Partition_of_India
    [pem_pos=0, final_scores=0.30707073,
final_local_scores=0.29726937, pem=-0.5217352,
similarity=15.7981]
```

```
32: 374012 Nawaz_Sharif
    [pem_pos=0, final_scores=0.8822592,
final_local_scores=0.87471396, pem=0.0, similarity=30.19292]
33: 193737 Benazir_Bhutto
    [pem_pos=0, final_scores=0.5378837,
final_local_scores=0.4939838, pem=0.0, similarity=18.831175]
34: 23235 Pakistan
    [pem_pos=0, final_scores=0.5487488,
final_local_scores=0.53283685, pem=-0.21944217,
similarity=21.183989]
35: 193737 Benazir_Bhutto
    [pem_pos=0, final_scores=0.53082037,
final_local_scores=0.48626816, pem=0.0, similarity=18.600925]
36: 193737 Benazir_Bhutto
    [pem_pos=0, final_scores=0.57229203,
final_local_scores=0.5315696, pem=0.0, similarity=19.95281]
```

Sample: b'746_India'

[1 1 India 1] ACC Apr-Jul '96 sales , output up . [2 2 BOMB
AY 2] 1996-08-28 [3 3 India 3] 's leading cement firm [4 4 As
sociated Cement Companies 4] (ACC) said on Wednesday its
cement sales rose to 3.1 million tonnes in April-July 1996
from 2.93 million a year ago . ACC Chairman [5 5 Nani Palkhiv
ala 5] told shareholders at the firm 's annual meeting cement
output rose to 3.14 million tonnes in the first quarter of
1996/97 (April-March) , from 3.01 million a year ago . [6 6
Palkhivala 6] said ACC had secured government approval to
take over a sick cement firm with a grinding capacity of
275,000 tonnes per year . " We will take it over early next
month , " he said . Talking about the cement industry in
general , [7 7 Palkhivala 7] said [8 8 Indian 8] production
rose by about 10 percent in 1995/96 . " The industry saw
capacity expansion of about 13 percent over 1994/95 from
77.79 million tonnes to 87.45 million tonnes , " [9 9 Palkhiv
ala 9] told shareholders . He said [10 10 Indian 10] cement
exports dropped about eight percent from the previous year
because of stiff international competition and inadequate
infrastructural facilities . ACC 's own export performance
was marginally better than in 1994/95 on account of a 36
percent rise in exports to [11 11 Nepal 11] and the opening
of a new market - [12 12 Sri Lanka 12] , he said . Despite
power shortages , ACC achieved a satisfactory growth in
production during the year with the help of its power
plants . ACC sold 9.4 million tonnes in 1995/96 , retaining
its top position in the [13 13 Indian 13] cement industry , [1
4 14 Palkhivala 14] said . -- [15 15 Bombay 15] newsroom
+91-22-265 9000

1: 14533 India
2: 19189 Mumbai
3: 14533 India
4: 2395785 ACC_Limited
5: 3500360 Nanabhoy_Palkhivala

```
6: 3500360 Nanabhoy_Palkhivala
7: 3500360 Nanabhoy_Palkhivala
8: 14533 India
9: 3500360 Nanabhoy_Palkhivala
10: 14533 India
11: 171166 Nepal
12: 26750 Sri_Lanka
13: 14533 India
14: 3500360 Nanabhoy_Palkhivala
15: 19189 Mumbai
1: 14533 India
    [pem_pos=0, final_scores=0.5378619,
final_local_scores=0.5070302, pem=-0.13622136,
similarity=19.961298]
2: 19189 Mumbai
    [pem_pos=0, final_scores=0.53315634,
final_local_scores=0.5291557, pem=-0.24257013,
similarity=21.199907]
3: 14533 India
    [pem_pos=0, final_scores=0.6446283,
final_local_scores=0.6236563, pem=-0.13622136,
similarity=23.441652]
4: 2395785 ACC_Limited
    [pem_pos=0, final_scores=0.49406046,
final_local_scores=0.6218635, pem=0.0, similarity=22.64736]
5: 3500360 Nanabhoy_Palkhivala
    [pem_pos=0, final_scores=0.5361218,
final_local_scores=0.574871, pem=0.0, similarity=21.24501]
6: 3500360 Nanabhoy_Palkhivala
    [pem_pos=0, final_scores=0.6342363,
final_local_scores=0.68204623, pem=0.0, similarity=24.44333]
7: 3500360 Nanabhoy_Palkhivala
    [pem_pos=0, final_scores=0.5771081,
final_local_scores=0.61964226, pem=0.0, similarity=22.581074]
8: 14533 India
    [pem_pos=0, final_scores=0.57750005,
final_local_scores=0.55032885, pem=-0.618294,
similarity=23.875004]
9: 3500360 Nanabhoy_Palkhivala
    [pem_pos=0, final_scores=0.5527081,
final_local_scores=0.5929891, pem=0.0, similarity=21.78569]
10: 14533 India
    [pem_pos=0, final_scores=0.6124902,
final_local_scores=0.58855027, pem=-0.618294,
similarity=25.015606]
11: 171166 Nepal
    [pem_pos=0, final_scores=0.58942646,
final_local_scores=0.6154996, pem=-0.22887385,
similarity=23.7021]
12: 26750 Sri_Lanka
    [pem_pos=0, final_scores=0.4505003,
final_local_scores=0.46000725, pem=-0.3027737,
```

```
similarity=19.463778]
13: 14533 India
    [pem_pos=0, final_scores=0.5909945,
final_local_scores=0.56506944, pem=-0.618294,
similarity=24.314892]
14: 3500360 Nanabhoy_Palkhivala
    [pem_pos=0, final_scores=0.58277863,
final_local_scores=0.6258365, pem=0.0, similarity=22.765923]
15: 19189 Mumbai
    [pem_pos=0, final_scores=0.3945455,
final_local_scores=0.37774444, pem=-0.24257013,
similarity=16.681494]
```

Sample: b'747_Belgium'

[1 1 Belgium 1] bank sanctions \$ 6.5 mln loan to [2 2 India 2] WSRL . [3 3 BOMBAY 3] 1996-08-28 [4 4 Belgium 4] 's [5 5 Kr edietbank 5] has approved a \$ 6.5 million loan to [6 6 India 6] 's Welspun Stahl Rohren Ltd (WSRL) to part-finance its submerged arc welded pipes plant , the [7 7 Indian 7] company said in a statement on Wednesday . The loan is at [8 8 LIBOR 8] plus one percent , it said . The loan , maturing in six years , is guaranteed by Indusind Bank for \$ 3.5 million and by [9 9 UTI Bank 9] for \$ 3 million , it said . The WSRL plant , located in the western [10 10 Indian 10] state of [11 11 Gujarat 11] , will have a capacity to manufacture 175,000 tonnes per annum of longitudinal pipes and 25,000 tonnes per annum of spiral welded pipes , the statement said . The longitudinal pipes plant is expected to be complete by the second quarter of September 1997 and the saw arc welded pipes plant by September 1996 , it said . WSRL is part of the Welspun group which has a presence in the cotton yarn , terry towels and polyester yarn industry , the statement said . -- [12 12 Bombay 12] newsroom +91-22-265 9000

```
1: 3343 Belgium
2: 14533 India
3: 19189 Mumbai
4: 3343 Belgium
5: 2181900 KBC_Bank
6: 14533 India
7: 14533 India
8: 236852 Libor
9: 12545216 Axis_Bank
10: 14533 India
11: 53707 Gujarat
12: 19189 Mumbai
1: 3343 Belgium
    [pem_pos=0, final_scores=0.46855575,
final_local_scores=0.49344593, pem=-0.24180195,
similarity=20.130081]
2: 14533 India
    [pem_pos=0, final_scores=0.73446757,
final_local_scores=0.70936996, pem=-0.13622136,
```

```
similarity=25.999521]
3: 19189 Mumbai
    [pem_pos=0, final_scores=0.5076543,
final_local_scores=0.48300087, pem=-0.24257013,
similarity=19.822556]
4: 3343 Belgium
    [pem_pos=0, final_scores=0.50538164,
final_local_scores=0.53367263, pem=-0.24180195,
similarity=21.330526]
5: 2181900 KBC_Bank
    [pem_pos=0, final_scores=0.31476617,
final_local_scores=0.4003122, pem=0.0, similarity=16.035828]
6: 14533 India
    [pem_pos=0, final_scores=0.8241094,
final_local_scores=0.8072901, pem=-0.13622136,
similarity=28.921652]
7: 14533 India
    [pem_pos=0, final_scores=0.5595176,
final_local_scores=0.5182637, pem=-0.618294,
similarity=22.918116]
8: 236852 Libor
    [pem_pos=0, final_scores=0.46001405,
final_local_scores=0.61853004, pem=0.0, similarity=22.547882]
9: 12545216 Axis_Bank
    [pem_pos=0, final_scores=0.45793146,
final_local_scores=0.5324951, pem=0.0, similarity=19.980429]
10: 14533 India
    [pem_pos=0, final_scores=0.5183392,
final_local_scores=0.47328258, pem=-0.618294,
similarity=21.575788]
11: 53707 Gujarat
    [pem_pos=0, final_scores=0.3050298,
final_local_scores=0.28511366, pem=0.0, similarity=12.598073]
12: 19189 Mumbai
    [pem_pos=0, final_scores=0.40819556,
final_local_scores=0.37435728, pem=-0.24257013,
similarity=16.580414]
```

Sample: b'748_India'

[1 1 India 1] fishermen say forced to carry Tamil refugees .
P.V. Krishnamoorthy [2 2 RAMESWARAM 2] , [3 3 India 3]
1996-08-28 [4 4 Indian 4] fishermen said on Wednesday they
had been forced at gunpoint to ferry refugees fleeing the
ethnic war in [5 5 Sri Lanka 5] to [6 6 India 6] , as a
protest strike by more than 30,000 fishermen entered its
ninth day . " There is little we can do when at mid-sea . The
[7 7 LTTE 7] ([8 8 Liberation Tigers of Tamil Eelam 8])
accosts us and , at the point of a gun , forces us to take
refugees , " said senior fishermen leader P. Arulanandam .
Some 850 refugees have landed in recent weeks at the port of [9 9 Rameswaram 9] in the southern [10 10 Indian 10] state of [11 11 Tamil Nadu 11] , home to 50 million Tamil-speaking

people , port officials say . [12 12 Rameswaram 12] is 15 km (10 miles) off the coast of [13 13 Sri Lanka 13] . State chief minister [14 14 M. Karunanidhi 14] has publicly welcomed the refugees , who are fleeing the 13-year war between Tamil separatists and government troops that [15 15 Colombo 15] says has cost 50,000 lives . But the influx has triggered fears of a repeat of the 1980s when some 200,000 refugees landed in [16 16 Tamil Nadu 16] . Intelligence officials said more than 5,000 [17 17 Tamils 17] were waiting on the western coast of [18 18 Sri Lanka 18] to cross into [19 19 India 19] . The [20 20 Indian 20] government has warned its fishermen that their boats would be impounded if they were caught ferrying refugees . The fishermen went on strike last week to protest the government 's cancellation of three trawlers ' fishing licenses after the boats were caught carrying Tamil refugees . All Fishermen 's Association secretary N.J. Bose said the strike would continue indefinitely and the fishermen would block road and rail traffic if their demands were not met . " Until the government releases our boats from naval custody and [21 21 Sri Lankan 21] naval custody , and gives us assurance (it will not revoke licences of boats ferrying refugees) , we will not call off our strike , " Bose said . [22 22 LTTE 22] spokesmen could not immediately be reached for comment , but [23 23 Sri Lankan 23] fishermen denied that [24 24 Indians 24] were being coerced into carrying refugees across the [25 25 Palk Strait 25] . " [26 26 Indian 26] fishermen come right up to [27 27 Pesalai 27] to fish , and when refugees request them to ferry them across , they readily oblige . Only some take money , " [28 28 Sri Lankan 28] boatman Chinnathambi said . Arulanandam denied the fishermen charged the refugees for passage to [29 29 India 29] and said it was unfair to penalise them for the refugees ' arrival . " We have not gone to sea since August 19 , but refugees are arriving daily nevertheless . How could we alone be held responsible for the influx ? " he said . An Indian Fisheries Department official said the government planned to urge fishermen not to cross the international boundary between [30 30 India 30] and [31 31 Sri Lanka 31] , but admitted this would be hard to enforce because of [32 32 Sri Lanka 32] 's pomfret-rich waters .

- 1: 14533 India
- 2: 373154 Rameswaram
- 3: 14533 India
- 4: 14533 India
- 5: 26750 Sri_Lanka
- 6: 14533 India
- 7: 18606 Liberation_Tigers_of_Tamil_Eelam
- 8: 18606 Liberation_Tigers_of_Tamil_Eelam
- 9: 373154 Rameswaram
- 10: 14533 India
- 11: 29918 Tamil_Nadu
- 12: 373154 Rameswaram

```
13: 26750 Sri_Lanka
14: 479912 Karunanidhi
15: 56636 Colombo
16: 29918 Tamil_Nadu
17: 300548 Tamil_people
18: 26750 Sri_Lanka
19: 14533 India
20: 14533 India
21: 26750 Sri_Lanka
22: 18606 Liberation_Tigers_of_Tamil_Eelam
23: 26750 Sri_Lanka
24: 14533 India
25: 373152 Palk_Strait
26: 14533 India
27: 12977962 Pesalai
28: 26750 Sri_Lanka
29: 14533 India
30: 14533 India
31: 26750 Sri_Lanka
32: 26750 Sri_Lanka
1: 14533 India
    [pem_pos=0, final_scores=0.67807966,
final_local_scores=0.6362287, pem=-0.13622136,
similarity=23.81684]
2: 373154 Rameswaram
    [pem_pos=0, final_scores=0.3468793,
final_local_scores=0.34963688, pem=0.0, similarity=14.523575]
3: 14533 India
    [pem_pos=0, final_scores=0.7544008,
final_local_scores=0.71959794, pem=-0.13622136,
similarity=26.304743]
4: 14533 India
    [pem_pos=0, final_scores=0.65750855,
final_local_scores=0.6137579, pem=-0.618294,
similarity=25.767853]
5: 26750 Sri_Lanka
    [pem_pos=0, final_scores=0.5249826,
final_local_scores=0.4770298, pem=-0.3027737,
similarity=19.971764]
6: 14533 India
    [pem_pos=0, final_scores=0.67998713,
final_local_scores=0.63831234, pem=-0.13622136,
similarity=23.879019]
7: 18606 Liberation_Tigers_of_Tamil_Eelam
    [pem_pos=0, final_scores=0.53394246,
final_local_scores=0.57527566, pem=0.0, similarity=21.257086]
8: 18606 Liberation_Tigers_of_Tamil_Eelam
    [pem_pos=0, final_scores=0.5449381,
final_local_scores=0.5872867, pem=0.0, similarity=21.61552]
9: 373154 Rameswaram
    [pem_pos=0, final_scores=0.38307673,
final_local_scores=0.38917708, pem=0.0, similarity=15.703533]
```

10: 14533 India
[pem_pos=0, final_scores=0.51745903,
final_local_scores=0.46077508, pem=-0.618294,
similarity=21.20254]
11: 29918 Tamil_Nadu
[pem_pos=0, final_scores=0.41678733,
final_local_scores=0.37917888, pem=0.0, similarity=15.405168]
12: 373154 Rameswaram
[pem_pos=0, final_scores=0.4285984,
final_local_scores=0.43890262, pem=0.0, similarity=17.187443]
13: 26750 Sri_Lanka
[pem_pos=0, final_scores=0.60350686,
final_local_scores=0.56280565, pem=-0.3027737,
similarity=22.531487]
14: 479912 Karunanidhi
[pem_pos=0, final_scores=0.501992,
final_local_scores=0.5730104, pem=0.0, similarity=21.189486]
15: 56636 Colombo
[pem_pos=0, final_scores=0.5124375,
final_local_scores=0.51283413, pem=-0.19393882,
similarity=20.448376]
16: 29918 Tamil_Nadu
[pem_pos=0, final_scores=0.38920373,
final_local_scores=0.349048, pem=0.0, similarity=14.506001]
17: 300548 Tamil_people
[pem_pos=0, final_scores=0.35292935,
final_local_scores=0.33306453, pem=-0.005057526,
similarity=14.056526]
18: 26750 Sri_Lanka
[pem_pos=0, final_scores=0.54270345,
final_local_scores=0.49638718, pem=-0.3027737,
similarity=20.549427]
19: 14533 India
[pem_pos=0, final_scores=0.59857136,
final_local_scores=0.5493779, pem=-0.13622136,
similarity=21.225039]
20: 14533 India
[pem_pos=0, final_scores=0.5939359,
final_local_scores=0.54431444, pem=-0.618294,
similarity=23.69552]
21: 26750 Sri_Lanka
[pem_pos=0, final_scores=0.5602793,
final_local_scores=0.5155861, pem=-0.27403015,
similarity=20.966047]
22: 18606 Liberation_Tigers_of_Tamil_Eelam
[pem_pos=0, final_scores=0.43326223,
final_local_scores=0.46529776, pem=0.0, similarity=17.975126]
23: 26750 Sri_Lanka
[pem_pos=0, final_scores=0.49858952,
final_local_scores=0.4481994, pem=-0.27403015,
similarity=18.955095]
24: 14533 India

```
[pem_pos=2, final_scores=0.4590066,
final_local_scores=0.39692467, pem=-1.8953385,
similarity=26.241884]
25: 373152 Palk_Strait
    [pem_pos=0, final_scores=0.45097774,
final_local_scores=0.49795848, pem=0.0, similarity=18.949787]
26: 14533 India
    [pem_pos=0, final_scores=0.44517112,
final_local_scores=0.3818115, pem=-0.618294,
similarity=18.84611]
27: 12977962 Pesalai
    [pem_pos=0, final_scores=0.4111004,
final_local_scores=0.53605705, pem=0.0, similarity=20.086725]
28: 26750 Sri_Lanka
    [pem_pos=0, final_scores=0.4193042,
final_local_scores=0.36159223, pem=-0.27403015,
similarity=16.370565]
29: 14533 India
    [pem_pos=0, final_scores=0.5746754,
final_local_scores=0.5232752, pem=-0.13622136,
similarity=20.446081]
30: 14533 India
    [pem_pos=0, final_scores=0.6044528,
final_local_scores=0.5558025, pem=-0.13622136,
similarity=21.416761]
31: 26750 Sri_Lanka
    [pem_pos=0, final_scores=0.4810918,
final_local_scores=0.4290858, pem=-0.3027737,
similarity=18.54102]
32: 26750 Sri_Lanka
    [pem_pos=0, final_scores=0.501322,
final_local_scores=0.45118415, pem=-0.3027737,
similarity=19.200478]
```

Sample: b'749_Indian'

[1 1 Indian 1] cotton trade shut for local festival . [2 2 B0
MBAY 2] 1996-08-28 [3 3 India 3] 's cotton trade was shut on
Wednesday for a local religious festival , dealers said .
Trading will resume on Thursday . On Tuesday , cotton prices
fell on profit-taking prompted by increased offerings from
state agencies . " Export deals remained thin and hardly a
few thousand bales were traded at the rate of 57.50 / 60
cents per pound , " one broker said . -- Bombay Commodities
+91-22-265 9000

```
1: 14533 India
2: 19189 Mumbai
3: 14533 India
1: 14533 India
    [pem_pos=0, final_scores=0.5424723,
final_local_scores=0.48516566, pem=-0.618294,
similarity=21.930403]
2: 19189 Mumbai
```

```
[pem_pos=0, final_scores=0.44753736,
final_local_scores=0.4413926, pem=-0.24257013,
similarity=18.580883]
3: 14533 India
[pem_pos=0, final_scores=0.645968,
final_local_scores=0.5982189, pem=-0.13622136,
similarity=22.682552]

Sample: b'750_VW'
VW sees group net profit doubling in Q3 . [1 1 DRESDEN 1] , [2 2 Germany 2] 1996-08-28 [3 3 German 3] carmaker [4 4 Volkswagen AG 4] said on Wednesday that it expected group net profit to double in the third quarter . " We have been seeking to double our profits ( during the period ) and we are confident of doing so , " VW chief financial officer Bruno Adelt told a briefing as part of the formal introduction of the new [5 5 VW Passat 5] sedan . Adelt did not give a concrete forecast . VW had a 1995 third quarter group net profit of 72 million marks . The group reported a 1996 first-half group net profit of 282 million marks . --
John Gilardi , Frankfurt Newsroom , +49 69 756525
1: 37410 Dresden
2: 11867 Germany
3: 11867 Germany
4: 32652 Volkswagen_Group
5: 554356 Volkswagen_Passat
1: 37410 Dresden
[pem_pos=0, final_scores=0.54005784,
final_local_scores=0.55820733, pem=-0.2008269,
similarity=21.83986]
2: 11867 Germany
[pem_pos=0, final_scores=0.8680318,
final_local_scores=0.87776476, pem=-0.15529114,
similarity=31.128464]
3: 11867 Germany
[pem_pos=1, final_scores=0.810891,
final_local_scores=0.815347, pem=-1.0465823,
similarity=34.112774]
4: 32652 Volkswagen_Group
[pem_pos=0, final_scores=0.16633825,
final_local_scores=0.19486797, pem=0.0, similarity=9.904962]
5: 554356 Volkswagen_Passat
[pem_pos=0, final_scores=0.34610158,
final_local_scores=0.4549302, pem=0.0, similarity=17.665737]
```

```
Sample: b'751_MILTIADIS'
[1 1 MILTIADIS EVERT 1] HEADS TO ALEXANDROUPOLIS THIS
WEEKEND . [2 2 ATHENS 2] 1996-08-28 Conservative New
Democracy ( ND ) party leader [3 3 Miltiadis Evert 3] will be
hitting the campaign trail and head to Alexandroupolis this
weekend to speak to the city 's businessmen on Sunday
morning , ND said . Evert will depart for Alexandroupolis on
```

Saturday afternoon . Prime Minister [4 4 Costas Simitis 4] criticized Evert today for unleashing a seven-point economic package that offers tax relief to merchants and other professionals , higher pensions to farmers and support for small business . The finance ministry estimated the cost of ND 's economic measures to over 600 billion drachmas but ND officials put the figure much lower to about 300 billion drachmas . [5 5 Simitis 5] blamed ND for the low absorption rate of EU funds and said the socialists will increase farmers ' pensions , combat tax evasion and accelerate GDP growth rates to 4.0 percent in a few years . Faster economic growth is a major component of ND 's economic programme and Evert has repeatedly blamed the socialists for the slow economic growth . -- Dimitris Kontogiannis , Athens Newsroom +301 3311812-4

```
1: 5481658 Miltiadis_Evert
2: 1216 Athens
3: 5481658 Miltiadis_Evert
4: 393824 Costas_Simitis
5: 393824 Costas_Simitis
1: 5481658 Miltiadis_Evert
    [pem_pos=0, final_scores=0.26288563,
final_local_scores=0.2671225, pem=0.0, similarity=12.061181]
2: 1216 Athens
    [pem_pos=0, final_scores=0.4082994,
final_local_scores=0.4999444, pem=-0.28831622,
similarity=20.57696]
3: 5481658 Miltiadis_Evert
    [pem_pos=0, final_scores=0.6256736,
final_local_scores=0.6634133, pem=0.0, similarity=23.88729]
4: 393824 Costas_Simitis
    [pem_pos=0, final_scores=0.6579295,
final_local_scores=0.68711776, pem=0.0, similarity=24.594677]
5: 393824 Costas_Simitis
    [pem_pos=0, final_scores=0.46712047,
final_local_scores=0.47868782, pem=0.0, similarity=18.374712]
```

Sample: b'752_Dow'

[1 1 Dow 1] pushes [2 2 London 2] stocks to new record . Peter Griffiths [3 3 LONDON 3] 1996-08-28 A firm start on [4 4 Wall Street 4] helped push leading 1 London shares to a new record high on Wednesday and [5 5 German 5] stocks closed floor trading up but the [6 6 Paris 6] bourse slipped sharply , hit by a weakening franc and fears of industrial unrest . On the foreign exchange markets , a survey indicating weaker than expected [7 7 Japanese 7] business sentiment boosted the dollar in early trading but it failed to build on its gains against the yen and slipped lower against the mark in quiet afternoon trade . 2 London shares , boosted by [8 8 Wall Street 8] , added to early gains with the blue chip [9 9 FTSE 100 9] index hitting a new peak of 3921.8 before dropping back slightly . One focus was [10 10 Br

itish Airways 10] , which rebounded after fears faded that the cancellation of " open skies " talks between the U.S. Transportation Department and the [11 11 British 11] government may jeopardise its tie-up with [12 12 American Air lines 12] . The conclusion of an open skies agreement had been made a prerequisite of the proposed link-up by the [13 13 U.S 13] . Better-than-expected [14 14 British 14] trade figures had little impact on equities . The non-EU July trade deficit totalled 506 million sterling (\$ 788 million) while June 's world deficit was 1.12 billion pounds . Forecasts were for deficits of 900 million sterling and 1.4 billion sterling . [15 15 German 15] shares shed some gains on profit-taking but nonetheless ended the floor session higher buoyed by demand for chemical stocks and a stable dollar . The 30-share [16 16 DAX 16] index closed at 2,563.16 points , up 4.32 . However , [17 17 French 17] stocks extended opening losses to more than one percent in morning trading , falling through the 2000 point resistance level on the main CAC-40 index . Dealers blamed a weakening franc and worries about the 1997 budget and possible autumn strikes Louis Viannet , leader of [18 18 France 18] 's Communist-led [19 19 CGT 19] union , criticised government plans for spending cuts in the 1997 budget on Wednesday and warned of labour unrest as [20 20 France 20] gets back to work after the holidays . Later in the day , helped by the firmer [21 21 Wall Street 21] opening , 3 Paris climbed back to the 2000 level but still remained well into negative territory . In foreign exchange the dollar was trading at around 108.40 yen in the [22 22 European 22] afternoon , up from its [23 23 European 23] close on Tuesday of 107.55 but off the day 's highs . It had been boosted overnight by the [24 24 Bank of Japan 24] 's quarterly corporate survey , or " tankan " , of major manufacturers -- an important gauge of business sentiment . The unexpectedly weak figures convinced markets that [25 25 Japanese 25] interest rates would stay at rock-bottom levels -- weakening the yen . The [26 26 Japanese 26] discount rate is currently at a record low of 0.5 percent . The [27 27 BOJ 27] sought to put the best face on the data which defied economists ' predictions of improving sentiment and was the first decline in business sentiment in a year . A [28 28 BOJ 28] spokesman said " the doors to the recovery have not been shut ... the recovery is still continuing . " These remarks gave the yen some backbone and the dollar failed to progress further . " The dollar 's gains were less than spectacular , suggesting that while it 's likely to benefit from revised [29 29 Japanese 29] rate expectations , a breakout from the established trading range does n't appear likely , " said currency economist Klaus Baader . Dealers said shares were also affected by a slippage in bond prices due to lower than expected June industrial production , showing the economy was still faltering . CURRENCIES AT 1500 [30 30 GMT 30] The dollar was at 108.40 yen and 1.4765 marks compared with

Tuesday 's [31 31 European 31] close of 107.78 yen and 1.4779 marks STOCK MARKETS AT 1500 [32 32 GMT 32] The Financial Times-Stock Exchange index of 100 leading [33 33 British 33] shares had risen 15 points to 3,920.7 . In [34 34 Paris 34] , the CAC-40 share index dropped 15.09 to 2,002.9 . The 30-share [35 35 DAX 35] index in [1 4 Frankfurt 1] closed up 4.32 at 2,563.16 . PRECIOUS METALS Gold fixed at \$ 388.50 versus Tuesday 's 5 London close of \$ 388.55 . Silver was at 521.15 cents . (\$ 1=.6421 Sterling)

```
1: 47361 Dow_Jones_Industrial_Average
2: 17867 London
3: 17867 London
4: 21560 New_York_Stock_Exchange
5: 11867 Germany
6: 22989 Paris
7: 15573 Japan
8: 21560 New_York_Stock_Exchange
9: 497717 FTSE_100_Index
10: 3970 British_Airways
11: 31717 United_Kingdom
12: 2386 American_Airlines
13: 3434750 United_States
14: 31717 United_Kingdom
15: 11867 Germany
16: 34153 DAX
17: 5843419 France
18: 5843419 France
19: 324642 General_Confederation_of_Labour_(France)
20: 5843419 France
21: 21560 New_York_Stock_Exchange
22: 9239 Europe
23: 9239 Europe
24: 188155 Bank_of_Japan
25: 15573 Japan
26: 15573 Japan
27: 188155 Bank_of_Japan
28: 188155 Bank_of_Japan
29: 15573 Japan
30: 12701 Greenwich_Mean_Time
31: 9239 Europe
32: 12701 Greenwich_Mean_Time
33: 31717 United_Kingdom
34: 22989 Paris
35: 34153 DAX
1: 197867 London_Stock_Exchange
    [winner: 17867 London, pem_pos=0,
final_scores=-0.15242465, final_local_scores=-0.23484921,
pem=-0.09086737, similarity=-2.424505]
    RECALL MISS: 197867 London_Stock_Exchange]
2: 197867 London_Stock_Exchange
    [winner: 17867 London, pem_pos=0,
final_scores=-0.23404051, final_local_scores=-0.32400215,
```

```
pem=-0.09086737, similarity=-5.0850058|
    RECALL MISS: 197867 London_Stock_Exchange]
3: 4694582 Euronext_Paris
    [winner: 22989 Paris, pem_pos=0,
final_scores=-0.41451937, final_local_scores=-0.50428855,
pem=-0.115301244, similarity=-10.332237|
    RECALL MISS: 4694582 Euronext_Paris]
4: 661174 Frankfurt_Stock_Exchange
    [winner: 10992 Frankfurt, pem_pos=0,
final_scores=0.08338319, final_local_scores=0.044216707,
pem=-0.07062824, similarity=5.793315|
    gt: 661174 Frankfurt_Stock_Exchange, pem_pos=6,
final_scores=-0.042160496, final_local_scores=-0.026231349,
pem=-3.349393, similarity=21.521437 ]
5: 197867 London_Stock_Exchange
    [winner: 17867 London, pem_pos=0,
final_scores=-0.29663908, final_local_scores=-0.39238158,
pem=-0.09086737, similarity=-7.125585|
    RECALL MISS: 197867 London_Stock_Exchange]
1: 47361 Dow_Jones_Industrial_Average
    [pem_pos=2, final_scores=0.46928442,
final_local_scores=0.54176813, pem=-1.5195237,
similarity=28.520561]
2: 17867 London
    [pem_pos=0, final_scores=0.46206343,
final_local_scores=0.44046527, pem=-0.09086737,
similarity=17.728226]
3: 17867 London
    [pem_pos=0, final_scores=0.7052433,
final_local_scores=0.7061023, pem=-0.09086737,
similarity=25.655365]
4: 21560 New_York_Stock_Exchange
    [pem_pos=7, final_scores=0.44214243,
final_local_scores=0.47668576, pem=-2.2964394,
similarity=30.803362]
5: 11867 Germany
    [pem_pos=1, final_scores=0.74661595,
final_local_scores=0.7519261, pem=-1.0465823,
similarity=32.220165]
6: 22989 Paris
    [pem_pos=0, final_scores=0.5150692,
final_local_scores=0.5162456, pem=-0.115301244,
similarity=20.122538]
7: 15573 Japan
    [pem_pos=0, final_scores=0.806403,
final_local_scores=0.8268191, pem=-0.99076414,
similarity=34.151573]
8: 21560 New_York_Stock_Exchange
    [pem_pos=7, final_scores=0.48805076,
final_local_scores=0.52683365, pem=-2.2964394,
similarity=32.299873]
9: 497717 FTSE_100_Index
```

```
[pem_pos=0, final_scores=0.6087554,
final_local_scores=0.6780383, pem=0.0, similarity=24.323729]
10: 3970 British_Airways
    [pem_pos=0, final_scores=0.43134242,
final_local_scores=0.4467544, pem=0.0, similarity=17.421755]
11: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.57449734,
final_local_scores=0.5474968, pem=-0.48936236,
similarity=23.089338]
12: 2386 American_Airlines
    [pem_pos=0, final_scores=0.39186096,
final_local_scores=0.44465792, pem=0.0, similarity=17.359192]
13: 3434750 United_States
    [pem_pos=0, final_scores=0.4598412,
final_local_scores=0.43366635, pem=0.0, similarity=17.031181]
14: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.58135206,
final_local_scores=0.5549845, pem=-0.48936236,
similarity=23.312786]
15: 11867 Germany
    [pem_pos=1, final_scores=0.7506471,
final_local_scores=0.7563296, pem=-1.0465823,
similarity=32.351574]
16: 34153 DAX
    [pem_pos=0, final_scores=0.68213654,
final_local_scores=0.7903717, pem=0.0, similarity=27.675982]
17: 5843419 France
    [pem_pos=0, final_scores=0.6376804,
final_local_scores=0.62633187, pem=-0.80223274,
similarity=27.143372]
18: 5843419 France
    [pem_pos=0, final_scores=0.6490037,
final_local_scores=0.6387008, pem=-0.16240834,
similarity=24.03302]
19: 324642 General_Confederation_of_Labour_(France)
    [pem_pos=1, final_scores=0.28738004,
final_local_scores=0.3538372, pem=-0.7709441,
similarity=18.841434]
20: 5843419 France
    [pem_pos=0, final_scores=0.6022845,
final_local_scores=0.5917588, pem=-0.16240834,
similarity=22.632177]
21: 21560 New_York_Stock_Exchange
    [pem_pos=7, final_scores=0.41268092,
final_local_scores=0.44450355, pem=-2.2964394,
similarity=29.842981]
22: 9239 Europe
    [pem_pos=0, final_scores=0.4432422,
final_local_scores=0.42528576, pem=-0.25557294,
similarity=18.170933]
23: 9239 Europe
    [pem_pos=0, final_scores=0.46134502,
```

```
final_local_scores=0.44506037, pem=-0.25557294,
similarity=18.761047]
24: 188155 Bank_of_Japan
    [pem_pos=0, final_scores=0.606659,
final_local_scores=0.70648015, pem=0.0, similarity=25.17249]
25: 15573 Japan
    [pem_pos=0, final_scores=0.7698782,
final_local_scores=0.7869213, pem=-0.99076414,
similarity=32.960945]
26: 15573 Japan
    [pem_pos=0, final_scores=0.7730537,
final_local_scores=0.79038996, pem=-0.99076414,
similarity=33.064453]
27: 188155 Bank_of_Japan
    [pem_pos=0, final_scores=0.5464187,
final_local_scores=0.6406768, pem=0.0, similarity=23.208786]
28: 188155 Bank_of_Japan
    [pem_pos=0, final_scores=0.5400467,
final_local_scores=0.63371634, pem=0.0, similarity=23.001072]
29: 15573 Japan
    [pem_pos=0, final_scores=0.7729576,
final_local_scores=0.79028517, pem=-0.99076414,
similarity=33.06133]
30: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.6269643,
final_local_scores=0.70345306, pem=0.0, similarity=25.082155]
31: 9239 Europe
    [pem_pos=0, final_scores=0.4202842,
final_local_scores=0.4002077, pem=-0.25557294,
similarity=17.422554]
32: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.5755399,
final_local_scores=0.64727974, pem=0.0, similarity=23.405832]
33: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.55854785,
final_local_scores=0.5300744, pem=-0.48936236,
similarity=22.56942]
34: 22989 Paris
    [pem_pos=0, final_scores=0.5863063,
final_local_scores=0.5940613, pem=-0.115301244,
similarity=22.444714]
35: 34153 DAX
    [pem_pos=0, final_scores=0.73073095,
final_local_scores=0.84345376, pem=0.0, similarity=29.260056]
1: 10992 Frankfurt
    [pem_pos=0, final_scores=0.08338319,
final_local_scores=0.044216707, pem=-0.07062824,
similarity=5.793315]
```

Sample: b'753_AEI'
AEI 's [1 1 Spanish 1] operation wins [2 2 ISO 9002 2] . [3 3
LONDON 3] 1996-08-28 Air Express International said in a

statement that [4 4 Spain 4] has become the twenty-second country in its network to achieve [5 5 ISO 9002 5] quality accreditation . It added [6 6 Bureau Veritas 6] has accredited AEI Iberfreight 's offices at [7 7 Alicante 7] , [8 8 Barcelona 8] , [9 9 Bilbao 9] , [10 10 Madrid 10] , [11 11 Seville 11] and [12 12 Valencia 12] as meeting the necessary standards . -- Air Cargo Newsroom Tel+44 171 542 7706 Fax+44 171 542 5017

1: 26667 Spain
2: 147042 ISO_9000
3: 17867 London
4: 26667 Spain
5: 147042 ISO_9000
6: 13946091 Bureau_Veritas
7: 2417 Alicante
8: 4443 Barcelona
9: 68029 Bilbao
10: 41188263 Madrid
11: 37770 Seville
12: 63861 Valencia
1: 26667 Spain
[pem_pos=1, final_scores=0.7110837, final_local_scores=0.70526737, pem=-1.2425895, similarity=31.893692]
2: 147042 ISO_9000
[pem_pos=0, final_scores=0.454432, final_local_scores=0.54788166, pem=0.0, similarity=20.439594]
3: 17867 London
[pem_pos=0, final_scores=0.7008619, final_local_scores=0.7258565, pem=-0.09086737, similarity=26.24487]
4: 26667 Spain
[pem_pos=0, final_scores=0.540795, final_local_scores=0.51925284, pem=-0.18857051, similarity=20.61073]
5: 147042 ISO_9000
[pem_pos=0, final_scores=0.50121135, final_local_scores=0.59898096, pem=0.0, similarity=21.9645]
6: 13946091 Bureau_Veritas
[pem_pos=0, final_scores=0.6270909, final_local_scores=0.7763133, pem=0.0, similarity=27.25645]
7: 2417 Alicante
[pem_pos=0, final_scores=0.36292446, final_local_scores=0.3289363, pem=-0.041829344, similarity=14.133303]
8: 4443 Barcelona
[pem_pos=0, final_scores=0.53564245, final_local_scores=0.5013018, pem=-0.37077636, similarity=21.065899]
9: 68029 Bilbao
[pem_pos=0, final_scores=0.46877402, final_local_scores=0.44409692, pem=0.0, similarity=17.342451]

```
10: 41188263 Madrid
    [pem_pos=0, final_scores=0.5624599,
final_local_scores=0.5219065, pem=-0.13236846,
similarity=20.384283]
11: 37770 Seville
    [pem_pos=0, final_scores=0.4777096,
final_local_scores=0.44563925, pem=-0.12539493,
similarity=18.070395]
12: 63861 Valencia
    [pem_pos=0, final_scores=0.43291014,
final_local_scores=0.39210063, pem=-0.9876697,
similarity=21.16188]
```

Sample: b'754_Arafat'

[1 1 Arafat 1] says [2 2 Israel 2] declares war on Palestinians . Wafa Amr [3 3 RAMALLAH 3] , West Bank 1996-08-28 [4 4 Palestinian 4] President [5 5 Yasser Arafat 5] said on Wednesday that [6 6 Israel 6] had declared war on the [7 7 Palestinians 7] and called for the first general strike in the West Bank and [8 8 Gaza 8] in two years . " What happened concerning continuous violations and crimes from this new [9 9 Israeli 9] leadership means they are declaring a state of war against the [10 10 Palestinian 10] people , " [11 11 Arafat 11] told the [12 12 Palestinian 12] legislature . Accusing [13 13 Israeli 13] Prime Minister [14 14 Benjamin Netanyahu 14] of stupidity , [15 15 Arafat 15] launched his strongest attack on the right-wing government since its election in May . The tirade was sparked by [16 16 Israel 16] 's announcement on Tuesday of plans to expand the [17 17 Jewish 17] settlement of Kiryat Sefer and its demolishing of a community centre in Arab East Jerusalem . " [18 18 Israel 18] has started the war on [19 19 Jerusalem 19] . They are idiots to have started the [20 20 Jerusalem 20] battle , " [21 21 Arafat 21] said in [22 22 Arabic 22] . " There will be no [23 23 Palestinian 23] state without [24 24 Jerusalem 24] . [25 25 Netanyahu 25] should know he is stupid to have started this battle . " [26 26 Arafat 26] called for a general strike " for [27 27 Jerusalem 27] " on Thursday in all of the West Bank and [28 28 Gaza Strip 28] . There has not been a joint shutdown there since May 1994 when [29 29 Israeli 29] troops began withdrawing under an interim self-rule agreement signed in 1993 . The strike will have little effect on the [30 30 Israeli 30] economy while hurting [31 31 Palestinian 31] merchants in [32 32 East Jerusalem 32] and [33 33 Bethlehem 33] who cater to the tourist trade . Some 25,000 [34 34 Palestinian 34] labourers are likely to stay away from their jobs , mainly in construction , in [35 35 Israel 35] . But [36 36 Palestinians 36] , once the backbone of the building industry , have been largely replaced by labourers from [37 37 Romania 37] and [1 1 China 1] . " On Friday , all [38 38 Moslems 38] , including [39 39 Palestinians 39] in [40 40 Israel 40] ... will go to ([41 41 Jerusalem 41] 's) [42 4

2 Al-Aqsa 42] mosque and pray . [43 43 Jews 43] and Christians who do not pray should accompany them and stand behind them , " [44 44 Arafat 44] said . Officials in [45 45 N etanyahu 45] 's office were not immediately available for comment on [46 46 Arafat 46] 's remarks . [47 47 Israeli 47] travel restrictions , imposed after bombings by [48 48 Moslem 48] militants in February and March , ban most of the two million [49 49 Arabs 49] in the West Bank and [50 50 Gaza 50] from entering [51 51 Jerusalem 51] . [52 52 Netanyahu 52] , who opposes trading occupied land for peace , made [53 53 Jerusalem 53] a central issue in his election campaign , accusing [54 54 Labour 54] Prime Minister [55 55 Shimon Peres 55] of planning to hand control of the Arab eastern part of the city to the [56 56 Palestinians 56] . [57 57 Israel 57] captured [58 58 East Jerusalem 58] in the 1967 [59 59 Middle East 59] war and claims both halves of the city as its capital . The PLO wants [60 60 East Jerusalem 60] as the capital of a future [61 61 Palestinian 61] state . In his speech to the crowded legislature , [62 62 Arafat 62] signalled he was not abandoning diplomacy although " alarm bells are ringing " . He unexpectedly broke off his address to take a telephone call from U.S. Middle East envoy [63 63 Dennis Ross 63] , who held talks in [64 64 Paris 64] on Tuesday with [65 65 Israeli 65] and [66 66 Egyptian 66] officials . When [67 67 Arafat 67] returned to the podium , he said senior PLO negotiator [68 68 Mahmoud Abbas 68] , better known as [69 69 Abu Mazen 69] , and [70 70 Netanyahu 70] aide [71 71 Dore Gold 71] could meet on Thursday . He quoted [72 72 Ross 72] as telling him : " The important thing is the [73 73 Israelis 73] are prepared to move . " [74 74 Arafat 74] said he replied : " is this like their previous promises ? " " No , they have good intentions , " [75 75 Arafat 75] quoted [76 76 Ross 76] as saying . [77 77 Palestinians 77] have been pressing [78 78 Israel 78] to carry out a long-delayed partial troop pullout from the flashpoint West Bank city of [79 79 Hebron 79] agreed by the previous [80 80 Labour 80] government . The [81 81 Israeli 81] government is studying redeployment plans submitted by Defence Minister [82 82 Yitzhak Mordechai 82] which are likely to demand a wider than agreed army presence in [83 83 Hebron 83] .

- 1: 34289 Yasser_Arafat
- 2: 9282173 Israel
- 3: 44132 Ramallah
- 4: 1114732 Palestine
- 5: 34289 Yasser_Arafat
- 6: 9282173 Israël
- 7: 1114732 Palestine
- 8: 55387 Gaza
- 9: 9282173 Israel
- 10: 1114732 Palestine
- 11: 34289 Yasser_Arafat
- 12: 1114732 Palestine

13: 9282173 Israel
14: 56472 Benjamin_Netanyahu
15: 34289 Yasser_Arafat
16: 9282173 Israel
17: 25955086 Jews
18: 9282173 Israel
19: 16043 Jerusalem
20: 16043 Jerusalem
21: 34289 Yasser_Arafat
22: 803 Arabic_language
23: 1114732 Palestine
24: 16043 Jerusalem
25: 56472 Benjamin_Netanyahu
26: 34289 Yasser_Arafat
27: 16043 Jerusalem
28: 12047 Gaza_Strip
29: 9282173 Israel
30: 9282173 Israel
31: 1114732 Palestine
32: 343508 East_Jerusalem
33: 4312 Bethlehem
34: 1114732 Palestine
35: 9282173 Israel
36: 1114732 Palestine
37: 25445 Romania
38: 6037917 Islam
39: 1114732 Palestine
40: 9282173 Israel
41: 16043 Jerusalem
42: 65759 Al-Aqsa_Mosque
43: 25955086 Jews
44: 34289 Yasser_Arafat
45: 56472 Benjamin_Netanyahu
46: 34289 Yasser_Arafat
47: 9282173 Israel
48: 6037917 Islam
49: 2185 Arab_people
50: 55387 Gaza
51: 16043 Jerusalem
52: 56472 Benjamin_Netanyahu
53: 16043 Jerusalem
54: 40312187 Israeli_Labor_Party
55: 48569 Shimon_Peres
56: 1114732 Palestine
57: 9282173 Israel
58: 343508 East_Jerusalem
59: 19323 Middle_East
60: 343508 East_Jerusalem
61: 1114732 Palestine
62: 34289 Yasser_Arafat
63: 1399016 Dennis_Ross
64: 22989 Paris

65: 9282173 Israel
66: 8087628 Egypt
67: 34289 Yasser_Arafat
68: 232595 Mahmoud_Abbas
69: 232595 Mahmoud_Abbas
70: 56472 Benjamin_Netanyahu
71: 1070942 Dore_Gold
72: 1399016 Dennis_Ross
73: 9282173 Israel
74: 34289 Yasser_Arafat
75: 34289 Yasser_Arafat
76: 1399016 Dennis_Ross
77: 1114732 Palestine
78: 9282173 Israel
79: 38577 Hebron
80: 40312187 Israeli_Labor_Party
81: 9282173 Israel
82: 1860312 Yitzhak_Mordechai
83: 38577 Hebron
1: 25734 Taiwan
 [winner: 5405 China, pem_pos=0, final_scores=0.40006584,
final_local_scores=0.45321947, pem=-0.15669827,
similarity=18.466835]
 gt: 25734 Taiwan, pem_pos=2, final_scores=-0.20005248,
final_local_scores=-0.20085132, pem=-4.365925,
similarity=21.838486]
1: 34289 Yasser_Arafat
 [pem_pos=0, final_scores=0.52557105,
final_local_scores=0.48417783, pem=0.0, similarity=18.538546]
2: 9282173 Israel
 [pem_pos=0, final_scores=0.6595504,
final_local_scores=0.6250891, pem=-0.21279564,
similarity=23.900833]
3: 44132 Ramallah
 [pem_pos=0, final_scores=0.460096,
final_local_scores=0.4253264, pem=0.0, similarity=16.7823]
4: 1114732 Palestine
 [pem_pos=1, final_scores=0.17478575,
final_local_scores=0.10263288, pem=-1.9100745,
similarity=17.539764]
5: 34289 Yasser_Arafat
 [pem_pos=0, final_scores=0.6116423,
final_local_scores=0.57734656, pem=0.0, similarity=21.318886]
6: 9282173 Israel
 [pem_pos=0, final_scores=0.6372969,
final_local_scores=0.60078055, pem=-0.21279564,
similarity=23.175419]
7: 1114732 Palestine
 [pem_pos=1, final_scores=0.3045919,
final_local_scores=0.24442637, pem=-1.6411982,
similarity=20.308973]
8: 55387 Gaza

```
[pem_pos=0, final_scores=0.29249328,
final_local_scores=0.23712152, pem=-0.30319342,
similarity=12.814706]
9: 9282173 Israel
    [pem_pos=1, final_scores=0.5628203,
final_local_scores=0.51942617, pem=-0.34790626,
similarity=21.482393]
10: 1114732 Palestine
    [pem_pos=1, final_scores=0.22906093,
final_local_scores=0.16192025, pem=-1.9100745,
similarity=19.309017]
11: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.62982124,
final_local_scores=0.5980553, pem=0.0, similarity=21.936876]
12: 1114732 Palestine
    [pem_pos=1, final_scores=0.32093996,
final_local_scores=0.26228416, pem=-1.9100745,
similarity=22.304077]
13: 9282173 Israel
    [pem_pos=0, final_scores=0.65863734,
final_local_scores=0.6240917, pem=-0.34790626,
similarity=24.605824]
14: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.70534176,
final_local_scores=0.6999414, pem=0.0, similarity=24.97736]
15: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.6313271,
final_local_scores=0.5997003, pem=0.0, similarity=21.985966]
16: 9282173 Israel
    [pem_pos=0, final_scores=0.56632286,
final_local_scores=0.5232522, pem=-0.21279564,
similarity=20.861818]
17: 25955086 Jews
    [pem_pos=1, final_scores=0.38278538,
final_local_scores=0.3704462, pem=-0.80476934,
similarity=19.521027]
18: 9282173 Israel
    [pem_pos=0, final_scores=0.606824,
final_local_scores=0.5674935, pem=-0.21279564,
similarity=22.182066]
19: 16043 Jerusalem
    [pem_pos=0, final_scores=0.6587966,
final_local_scores=0.63520175, pem=-0.13239029,
similarity=23.76536]
20: 16043 Jerusalem
    [pem_pos=0, final_scores=0.5930061,
final_local_scores=0.5633356, pem=-0.13239029,
similarity=21.62073]
21: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.6553282,
final_local_scores=0.62591785, pem=0.0, similarity=22.76835]
22: 803 Arabic_language
```

```
[pem_pos=0, final_scores=0.35969853,
final_local_scores=0.38215947, pem=-0.084880665,
similarity=15.955708]
23: 1114732 Palestine
    [pem_pos=1, final_scores=0.26080376,
final_local_scores=0.19659448, pem=-1.9100745,
similarity=20.343765]
24: 16043 Jerusalem
    [pem_pos=0, final_scores=0.52420276,
final_local_scores=0.48817837, pem=-0.13239029,
similarity=19.377888]
25: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.5358283,
final_local_scores=0.5160497, pem=0.0, similarity=19.489664]
26: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.61370057,
final_local_scores=0.580446, pem=0.0, similarity=21.411379]
27: 16043 Jerusalem
    [pem_pos=0, final_scores=0.5426941,
final_local_scores=0.5083774, pem=-0.13239029,
similarity=19.980667]
28: 12047 Gaza_Strip
    [pem_pos=0, final_scores=0.41588616,
final_local_scores=0.38328445, pem=0.0, similarity=15.527686]
29: 9282173 Israel
    [pem_pos=1, final_scores=0.54164165,
final_local_scores=0.4962917, pem=-0.34790626,
similarity=20.792015]
30: 9282173 Israel
    [pem_pos=1, final_scores=0.5480781,
final_local_scores=0.50332254, pem=-0.34790626,
similarity=21.001831]
31: 1114732 Palestine
    [pem_pos=1, final_scores=0.34702915,
final_local_scores=0.29078257, pem=-1.9100745,
similarity=23.154526]
32: 343508 East_Jerusalem
    [pem_pos=0, final_scores=0.37244052,
final_local_scores=0.34753543, pem=0.0, similarity=14.460863]
33: 4312 Bethlehem
    [pem_pos=0, final_scores=0.41374135,
final_local_scores=0.3968823, pem=-0.5095743,
similarity=18.704617]
34: 1114732 Palestine
    [pem_pos=1, final_scores=0.31420875,
final_local_scores=0.25493133, pem=-1.9100745,
similarity=22.084654]
35: 9282173 Israel
    [pem_pos=0, final_scores=0.5418966,
final_local_scores=0.49657017, pem=-0.21279564,
similarity=20.065573]
36: 1114732 Palestine
```

[pem_pos=1, final_scores=0.3466435,
final_local_scores=0.2903614, pem=-1.6411982,
similarity=21.679768]
37: 25445 Romania
[pem_pos=0, final_scores=0.653008,
final_local_scores=0.73433715, pem=-0.21241479,
similarity=27.158943]
38: 6037917 Islam
[pem_pos=0, final_scores=0.50347215,
final_local_scores=0.5384877, pem=0.0, similarity=20.15926]
39: 1114732 Palestine
[pem_pos=1, final_scores=0.40147954,
final_local_scores=0.3502614, pem=-1.6411982,
similarity=23.467302]
40: 9282173 Israel
[pem_pos=0, final_scores=0.587592,
final_local_scores=0.5464855, pem=-0.21279564,
similarity=21.555147]
41: 16043 Jerusalem
[pem_pos=0, final_scores=0.57145655,
final_local_scores=0.539796, pem=-0.13239029,
similarity=20.91826]
42: 65759 Al-Aqsa_Mosque
[pem_pos=0, final_scores=0.4212618,
final_local_scores=0.45205867, pem=0.0, similarity=17.580046]
43: 25955086 Jews
[pem_pos=0, final_scores=0.41765487,
final_local_scores=0.40853578, pem=-0.1252217,
similarity=16.96221]
44: 34289 Yasser_Arafat
[pem_pos=0, final_scores=0.6395846,
final_local_scores=0.6087203, pem=0.0, similarity=22.25514]
45: 56472 Benjamin_Netanyahu
[pem_pos=0, final_scores=0.49929184,
final_local_scores=0.47613925, pem=0.0, similarity=18.298658]
46: 34289 Yasser_Arafat
[pem_pos=0, final_scores=0.63158077,
final_local_scores=0.59997743, pem=0.0, similarity=21.994236]
47: 9282173 Israel
[pem_pos=1, final_scores=0.494924,
final_local_scores=0.44525975, pem=-0.34790626,
similarity=19.26912]
48: 6037917 Islam
[pem_pos=1, final_scores=0.44654363,
final_local_scores=0.4763019, pem=-1.2150909,
similarity=24.911364]
49: 2185 Arab_people
[pem_pos=1, final_scores=0.43889302,
final_local_scores=0.43874818, pem=-1.8745545,
similarity=27.376953]
50: 55387 Gaza
[pem_pos=0, final_scores=0.334257,

```
final_local_scores=0.28274202, pem=-0.30319342,
similarity=14.176111]
51: 16043 Jerusalem
    [pem_pos=0, final_scores=0.5886752,
final_local_scores=0.5586047, pem=-0.13239029,
similarity=21.479551]
52: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.50623363,
final_local_scores=0.4837221, pem=0.0, similarity=18.524944]
53: 16043 Jerusalem
    [pem_pos=0, final_scores=0.5597788,
final_local_scores=0.5270398, pem=-0.13239029,
similarity=20.537592]
54: 40312187 Israeli_Labor_Party
    [pem_pos=7, final_scores=0.46074134,
final_local_scores=0.47812635, pem=-3.1950269,
similarity=35.73301]
55: 48569 Shimon_Peres
    [pem_pos=0, final_scores=0.6358219,
final_local_scores=0.63835144, pem=0.0, similarity=23.139393]
56: 1114732 Palestine
    [pem_pos=1, final_scores=0.3536151,
final_local_scores=0.2979768, pem=-1.6411982,
similarity=21.907026]
57: 9282173 Israel
    [pem_pos=0, final_scores=0.5254744,
final_local_scores=0.47863144, pem=-0.21279564,
similarity=19.530247]
58: 343508 East_Jerusalem
    [pem_pos=0, final_scores=0.34619957,
final_local_scores=0.31887123, pem=0.0, similarity=13.605467]
59: 19323 Middle_East
    [pem_pos=1, final_scores=0.40615076,
final_local_scores=0.40251118, pem=0.0, similarity=16.10145]
60: 343508 East_Jerusalem
    [pem_pos=0, final_scores=0.38803577,
final_local_scores=0.36457092, pem=0.0, similarity=14.969235]
61: 1114732 Palestine
    [pem_pos=1, final_scores=0.33953762,
final_local_scores=0.28259927, pem=-1.9100745,
similarity=22.91032]
62: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.5861773,
final_local_scores=0.55038095, pem=0.0, similarity=20.514177]
63: 1399016 Dennis_Ross
    [pem_pos=0, final_scores=0.5541703,
final_local_scores=0.60276335, pem=0.0, similarity=22.077374]
64: 22989 Paris
    [pem_pos=0, final_scores=0.59970677,
final_local_scores=0.6696725, pem=-0.115301244,
similarity=24.701101]
65: 9282173 Israel
```

[pem_pos=1, final_scores=0.55316955,
final_local_scores=0.50888413, pem=-0.34790626,
similarity=21.167799]
66: 8087628 Egypt
[pem_pos=0, final_scores=0.4056903,
final_local_scores=0.37966937, pem=-1.052751,
similarity=21.144827]
67: 34289 Yasser_Arafat
[pem_pos=0, final_scores=0.5626803,
final_local_scores=0.5247141, pem=0.0, similarity=19.748228]
68: 232595 Mahmoud_Abbas
[pem_pos=0, final_scores=0.55134183,
final_local_scores=0.5463944, pem=0.0, similarity=20.395212]
69: 232595 Mahmoud_Abbas
[pem_pos=0, final_scores=0.65478116,
final_local_scores=0.65938616, pem=0.0, similarity=23.76711]
70: 56472 Benjamin_Netanyahu
[pem_pos=0, final_scores=0.50296587,
final_local_scores=0.48015255, pem=0.0, similarity=18.418423]
71: 1070942 Dore_Gold
[pem_pos=0, final_scores=0.62552285,
final_local_scores=0.66641796, pem=0.0, similarity=23.976954]
72: 1399016 Dennis_Ross
[pem_pos=0, final_scores=0.77379805,
final_local_scores=0.8426733, pem=0.0, similarity=29.236765]
73: 9282173 Israel
[pem_pos=1, final_scores=0.47270292,
final_local_scores=0.4209866, pem=-0.5460739,
similarity=19.622427]
74: 34289 Yasser_Arafat
[pem_pos=0, final_scores=0.61175925,
final_local_scores=0.5783254, pem=0.0, similarity=21.348097]
75: 34289 Yasser_Arafat
[pem_pos=0, final_scores=0.6416778,
final_local_scores=0.61100686, pem=0.0, similarity=22.323376]
76: 1399016 Dennis_Ross
[pem_pos=0, final_scores=0.74772006,
final_local_scores=0.814187, pem=0.0, similarity=28.386677]
77: 1114732 Palestine
[pem_pos=1, final_scores=0.2235081,
final_local_scores=0.15585464, pem=-1.6411982,
similarity=17.665817]
78: 9282173 Israel
[pem_pos=0, final_scores=0.64013535,
final_local_scores=0.60388106, pem=-0.21279564,
similarity=23.267944]
79: 38577 Hebron
[pem_pos=0, final_scores=0.4250998,
final_local_scores=0.4066077, pem=-0.39021417,
similarity=18.345743]
80: 40312187 Israeli_Labor_Party
[pem_pos=7, final_scores=0.46468216,

```
final_local_scores=0.4824311, pem=-3.1950269,
similarity=35.861473]
81: 9282173 Israel
    [pem_pos=1, final_scores=0.5671304,
final_local_scores=0.5241342, pem=-0.34790626,
similarity=21.62289]
82: 1860312 Yitzhak_Mordechai
    [pem_pos=0, final_scores=0.40763634,
final_local_scores=0.4764195, pem=0.0, similarity=18.307022]
83: 38577 Hebron
    [pem_pos=0, final_scores=0.40630817,
final_local_scores=0.38608074, pem=-0.39021417,
similarity=17.733177]
1: 5405 China
    [pem_pos=0, final_scores=0.40006584,
final_local_scores=0.45321947, pem=-0.15669827,
similarity=18.466835]
```

Sample: b'755_Turkey'

Turkey 's [1 1 Ciller 1] to hold talks in [2 2 Jordan 2] . [3 3 ANKARA 3] 1996-08-28 Turkish Foreign Minister [4 4 Tansu Ciller 4] will visit [5 5 Jordan 5] on September 3 on her first trip abroad since she took office in June , [6 6 Foreign Ministry 6] spokesman Omer Akbel said on Wednesday . " The two-day visit will take place at the invitation of [7 7 Jordanian 7] Prime Minister Abdul-Karim al-Kabariti , " he told reporters . He said Turkey considered [8 8 Jordan 8] a suitable country with which to cooperate on [9 9 Middle East 9] matters . Bilateral relations and the [10 10 Middle East 10] peace process would be on the table during the visit , Akbel said . Turkish Prime Minister [11 11 Necmettin Erbakan 11] visited mainly [12 12 Moslem 12] countries , including [13 13 Iran 13] , during a 10-day tour earlier in August .

```
1: 753267 Tansu_Çiller
2: 7515964 Jordan
3: 802 Ankara
4: 753267 Tansu_Çiller
5: 7515964 Jordan
6: 10283896 Ministry_of_Foreign_Affairs_(Turkey)
7: 7515964 Jordan
8: 7515964 Jordan
9: 19323 Middle_East
10: 19323 Middle_East
11: 1328699 Necmettin_Erbakan
12: 6037917 Islam
13: 14653 Iran
1: 753267 Tansu_Çiller
    [pem_pos=0, final_scores=0.6181216,
final_local_scores=0.6895727, pem=0.0, similarity=24.667938]
2: 7515964 Jordan
    [pem_pos=0, final_scores=0.5313682,
final_local_scores=0.493047, pem=-0.53647494,
```

```
similarity=21.720654]
3: 802 Ankara
    [pem_pos=0, final_scores=0.41641933,
final_local_scores=0.41624242, pem=0.0, similarity=16.511217]
4: 753267 Tansu Çiller
    [pem_pos=0, final_scores=0.7647241,
final_local_scores=0.84971356, pem=0.0, similarity=29.446861]
5: 7515964 Jordan
    [pem_pos=0, final_scores=0.6041236,
final_local_scores=0.57252127, pem=-0.53647494,
similarity=24.092323]
6: 10283896 Ministry_of_Foreign_Affairs_(Turkey)
    [pem_pos=9, final_scores=0.19428568,
final_local_scores=0.23671699, pem=-2.9282627,
similarity=27.078167]
7: 7515964 Jordan
    [pem_pos=0, final_scores=0.55881315,
final_local_scores=0.52302647, pem=-0.18608934,
similarity=20.709848]
8: 7515964 Jordan
    [pem_pos=0, final_scores=0.54693526,
final_local_scores=0.51005167, pem=-0.53647494,
similarity=22.228106]
9: 19323 Middle_East
    [pem_pos=0, final_scores=0.417697,
final_local_scores=0.3923385, pem=0.0, similarity=15.797876]
10: 19323 Middle_East
    [pem_pos=0, final_scores=0.4126892,
final_local_scores=0.38686824, pem=0.0, similarity=15.634632]
11: 1328699 Necmettin_Erbakan
    [pem_pos=0, final_scores=0.60901314,
final_local_scores=0.68664056, pem=0.0, similarity=24.580437]
12: 6037917 Islam
    [pem_pos=1, final_scores=0.4380942,
final_local_scores=0.46607965, pem=-1.2150909,
similarity=24.606312]
13: 14653 Iran
    [pem_pos=0, final_scores=0.49982685,
final_local_scores=0.5663063, pem=-0.13856095,
similarity=21.742939]
```

Sample: b'756_Italian'

```
[1 1 Italian 1] prime minister to visit Turkey . [2 2 ANKARA
2] 1996-08-28 [3 3 Italian 3] Prime Minister [4 4 Romano Prodi
4] will pay a one-day working visit to Turkey on September
3 , the Turkish [5 5 Foreign Ministry 5] said on Wednesday .
" Both countries ' governments were formed recently . This
visit will create a direct contact opportunity for the two
parties to express their views , " spokesman Omer Akbel told
a news briefing . The [6 6 Italian 6] prime minister , in
office since May 18 , will be the first [7 7 Western 7]
leader to meet [8 8 Islamist 8] Prime Minister [9 9 Necmettin
```

Erbakan 9] since he came to power on June 28 . [10 10 Prodi 10] will also meet President [11 11 Suleyman Demirel 11] and Foreign Minister [12 12 Tansu Ciller 12] , Akbel said .

1: 14532 Italy
2: 802 Ankara
3: 14532 Italy
4: 197983 Romano_Prodi
5: 10283896 Ministry_of_Foreign_Affairs_(Turkey)
6: 14532 Italy
7: 21208200 Western_world
8: 6037917 Islam
9: 1328699 Necmettin_Erbakan
10: 197983 Romano_Prodi
11: 166093 Süleyman_Demirel
12: 753267 Tansu_Çiller

1: 14532 Italy
[pem_pos=0, final_scores=0.634386,
final_local_scores=0.66223854, pem=-0.75948656,
similarity=27.982437]
2: 802 Ankara
[pem_pos=0, final_scores=0.41904277,
final_local_scores=0.42987227, pem=0.0, similarity=16.91796]
3: 14532 Italy
[pem_pos=0, final_scores=0.72985464,
final_local_scores=0.76652354, pem=-0.75948656,
similarity=31.094511]
4: 197983 Romano_Prodi
[pem_pos=0, final_scores=0.65401083,
final_local_scores=0.69569916, pem=0.0, similarity=24.850763]
5: 10283896 Ministry_of_Foreign_Affairs_(Turkey)
[pem_pos=9, final_scores=0.20974974,
final_local_scores=0.26024866, pem=-2.9282627,
similarity=27.7804]
6: 14532 Italy
[pem_pos=0, final_scores=0.680325,
final_local_scores=0.7124199, pem=-0.75948656,
similarity=29.47995]
7: 21208200 Western_world
[pem_pos=0, final_scores=0.42194408,
final_local_scores=0.4437927, pem=-1.5917408,
similarity=25.989506]
8: 6037917 Islam
[pem_pos=1, final_scores=0.31368405,
final_local_scores=0.35783187, pem=-1.3455799,
similarity=22.085602]
9: 1328699 Necmettin_Erbakan
[pem_pos=0, final_scores=0.57932216,
final_local_scores=0.6368117, pem=0.0, similarity=23.093443]
10: 197983 Romano_Prodi
[pem_pos=0, final_scores=0.3153817,
final_local_scores=0.32579827, pem=0.0, similarity=13.812183]
11: 166093 Süleyman_Demirel

```
[pem_pos=0, final_scores=0.629876,
final_local_scores=0.6569166, pem=0.0, similarity=23.693415]
12: 753267 Tansu Çiller
[pem_pos=0, final_scores=0.68970937,
final_local_scores=0.76094055, pem=0.0, similarity=26.797697]

Sample: b'757_Turkey'
Turkey says 25 [1 1 Kurd 1] rebels killed in clashes . [1
DIYARBAKIR 1] , Turkey 1996-08-28 Turkish troops killed 25 [2
2 Kurdish 2] rebels in recent clashes in the east of the
country , security officials said on Wednesday . The
emergency rule governor 's office said in a statement that 10
rebels from the [3 3 Kurdistan Workers Party 3] ( [4 4 PKK 4]
) were killed in fighting in [5 5 Tunceli 5] province .
Soldiers killed nine more [6 6 PKK 6] guerrillas in [7 7 Sırnak
ak 7] province and six in [2 1 Hakkari 2] . The statement did
not report any military casualties and did not say exactly
when the clashes took place . Over 20,000 people have been
killed in the [8 8 PKK 8] 's fight for independence or
autonomy in southeastern Turkey .
1: 80777 Kurdistan
2: 80777 Kurdistan
3: 69680 Kurdistan_Workers'_Party
4: 69680 Kurdistan_Workers'_Party
5: 291275 Tunceli_Province
6: 69680 Kurdistan_Workers'_Party
7: 10541472 Şırnak_Province
8: 69680 Kurdistan_Workers'_Party
1: 2463135 Hakkâri
[winner: 1372482 Hakkâri_Province, pem_pos=1,
final_scores=0.100296214, final_local_scores=0.102442145,
pem=-0.47937384, similarity=9.7537]
gt: 2463135 Hakkâri, pem_pos=0,
final_scores=0.060376838, final_local_scores=0.06725834,
pem=0.0, similarity=6.096836 ]
1: 80777 Kurdistan
[pem_pos=3, final_scores=0.2539472,
final_local_scores=0.22295058, pem=-2.608398,
similarity=24.927876]
2: 80777 Kurdistan
[pem_pos=9, final_scores=0.21450083,
final_local_scores=0.17986143, pem=-2.8001814,
similarity=24.684958]
3: 69680 Kurdistan_Workers'_Party
[pem_pos=0, final_scores=0.6569679,
final_local_scores=0.6364653, pem=0.0, similarity=23.083107]
4: 69680 Kurdistan_Workers'_Party
[pem_pos=0, final_scores=0.57358867,
final_local_scores=0.54538614, pem=0.0, similarity=20.365124]
5: 291275 Tunceli_Province
[pem_pos=0, final_scores=0.3199249,
final_local_scores=0.3131428, pem=0.0, similarity=13.434519]
```

```
6: 69680 Kurdistan_Workers'_Party
    [pem_pos=0, final_scores=0.57611006,
final_local_scores=0.5481404, pem=0.0, similarity=20.447315]
7: 10541472 Şırnak_Province
    [pem_pos=1, final_scores=0.38558376,
final_local_scores=0.43257427, pem=-0.68484837,
similarity=20.722904]
8: 69680 Kurdistan_Workers'_Party
    [pem_pos=0, final_scores=0.5406461,
final_local_scores=0.5094014, pem=0.0, similarity=19.291265]
1: 642286 Diyarbakır
    [pem_pos=0, final_scores=0.053852335,
final_local_scores=0.013808042, pem=-0.24235578,
similarity=5.819741]
2: 1372482 Hakkâri_Province
    [pem_pos=1, final_scores=0.100296214,
final_local_scores=0.102442145, pem=-0.47937384,
similarity=9.7537]
```

Sample: b'758_Israel'

[1 1 Israel 1] 's Levy to meet [2 2 Mubarak 2] in [3 3 Egypt 3] . [4 4 JERUSALEM 4] 1996-08-28 [5 5 Israeli 5] Foreign Minister [6 6 David Levy 6] will visit [7 7 Egypt 7] this Sunday for talks with President [8 8 Hosni Mubarak 8] , the [9 9 Foreign Ministry 9] said on Wednesday . The trip will be Levy 's first to an Arab state as a minister in Prime Minister [10 10 Benjamin Netanyahu 10] 's government .

```
1: 9282173 Israel
2: 51884 Hosni_Mubarak
3: 8087628 Egypt
4: 16043 Jerusalem
5: 9282173 Israel
6: 1256116 David_Levy_(Israeli_politician)
7: 8087628 Egypt
8: 51884 Hosni_Mubarak
9: 18133296 Ministry_of_Foreign_Affairs_(Israel)
10: 56472 Benjamin_Netanyahu
1: 9282173 Israel
    [pem_pos=0, final_scores=0.5324438,
final_local_scores=0.50113827, pem=-0.21279564,
similarity=20.201895]
2: 51884 Hosni_Mubarak
    [pem_pos=0, final_scores=0.48775166,
final_local_scores=0.4792453, pem=0.0, similarity=18.391348]
3: 8087628 Egypt
    [pem_pos=0, final_scores=0.50490594,
final_local_scores=0.493877, pem=-0.27480945,
similarity=20.322445]
4: 16043 Jerusalem
    [pem_pos=0, final_scores=0.5215479,
final_local_scores=0.51650727, pem=-0.13239029,
similarity=20.223278]
```

```
5: 9282173 Israel
    [pem_pos=0, final_scores=0.60398144,
final_local_scores=0.5792822, pem=-0.34790626,
similarity=23.268618]
6: 1256116 David_Levy_(Israeli_politician)
    [pem_pos=1, final_scores=0.5828703,
final_local_scores=0.6173658, pem=-1.1434275,
similarity=28.731276]
7: 8087628 Egypt
    [pem_pos=0, final_scores=0.52278906,
final_local_scores=0.5134116, pem=-0.27480945,
similarity=20.905396]
8: 51884 Hosni_Mubarak
    [pem_pos=0, final_scores=0.59063596,
final_local_scores=0.5777111, pem=0.0, similarity=21.329765]
9: 18133296 Ministry_of_Foreign_Affairs_(Israel)
    [pem_pos=3, final_scores=0.10058318,
final_local_scores=0.06256455, pem=-2.603179,
similarity=20.113256]
10: 56472 Benjamin_Netanyahu
    [pem_pos=0, final_scores=0.6596771,
final_local_scores=0.6560217, pem=0.0, similarity=23.66671]
```

Sample: b'759_FSA'

FSA qualifies five muni bond issues for insurance . [1 1 NEW YORK 1] 1996-08-28 Financial Security Assurance said Wednesday it qualified for bond insurance the following five municipal issues scheduled for competitive sale today : -- 1 P aris School District No 7 , Ark . , \$ 2.44 million refunding bonds . -- St Ansgar Community School District , Iowa , \$ 3.334 million general obligation school bonds . -- [2 2 Avalon Borough 2] , [3 3 N.J. 3] , \$ 11.4 million GOs . -- Seaford Union Free School District , [4 4 N.Y. 4] , \$ 5 million school bonds . -- [5 5 Akron 5] , [6 6 Ohio 6] , \$ 6.31 million improvement bonds . -- U.S. Municipal Desk , 212-859-1650

```
1: 645042 New_York_City
2: 125107 Avalon,_New_Jersey
3: 21648 New_Jersey
4: 8210131 New_York
5: 129842 Akron,_Ohio
6: 22199 Ohio
1: 107104 Paris,_Arkansas
    [winner: 38579741 Paris_School_District, pem_pos=0,
final_scores=-0.387949, final_local_scores=-0.2940917,
pem=0.0, similarity=-4.6865706|
     RECALL MISS: 107104 Paris,_Arkansas]
1: 645042 New_York_City
    [pem_pos=1, final_scores=0.4883269,
final_local_scores=0.43700874, pem=-1.7906687,
similarity=26.86886]
2: 125107 Avalon,_New_Jersey
```

```
[pem_pos=0, final_scores=0.4639737,
final_local_scores=0.53215396, pem=0.0, similarity=19.97025]
3: 21648 New_Jersey
    [pem_pos=0, final_scores=0.53305507,
final_local_scores=0.48447573, pem=0.0, similarity=18.547436]
4: 8210131 New_York
    [pem_pos=0, final_scores=0.42711073,
final_local_scores=0.37417132, pem=0.0, similarity=15.255731]
5: 129842 Akron,_Ohio
    [pem_pos=0, final_scores=0.4604463,
final_local_scores=0.44594753, pem=-0.6020413,
similarity=20.67167]
6: 22199 Ohio
    [pem_pos=0, final_scores=0.66199,
final_local_scores=0.6413318, pem=-0.19147073,
similarity=24.26958]
```

Sample: b'760_Colo'

```
1 Colo . taxable health deal rated Aa2 / VMIG-1 - Moody 's .
[1 1 NEW YORK 1] 1996-08-28 Moody 's Investors Service -
Rating Announcement As of 08/26/96 . Issuer : Colorado Health
Fac . Auth . National Benevolent Assoc . - Colorado Christian
Home Proj . Series 1996 B Taxable State : CO Rating : Aa2 /
VMIG 1 Sale Amount : 4,300,000 Expected Sale Date : 08/26/96
1: 645042 New_York_City
1: 5399 Colorado
    [winner: 114241 Colo,_Iowa, pem_pos=0,
final_scores=-0.30467784, final_local_scores=-0.20271987,
pem=-0.522792, similarity=0.88317215]
    gt: 5399 Colorado, pem_pos=13, final_scores=-0.6609121,
final_local_scores=-0.7262425, pem=-4.919981,
similarity=9.172811 ]
    [winner: 4721653 Zyphe_language, pem_pos=12,
final_scores=-1.3050001, final_local_scores=-1.2218909,
pem=-4.4188404, similarity=-8.343606]
    RECALL MISS: 5399 Colorado]
1: 645042 New_York_City
    [pem_pos=1, final_scores=0.165855,
final_local_scores=0.4537586, pem=-1.7906687,
similarity=27.36871]
```

Sample: b'761_Lamm'

```
[1 1 Lamm 1] will not endorse [2 2 Perot 2] for Reform
ticket-CNN . [3 3 WASHINGTON 3] 1996-08-27 Former [4 4 Colora
do 4] [5 5 Democratic 5] Gov . [6 6 Richard Lamm 6] has
decided not to endorse [7 7 Ross Perot 7] as the presidential
candidate of the [8 8 Reform Party 8] , [9 9 CNN 9] reported
late Tuesday . [10 10 CNN 10] quoted aides and family members
as saying [11 11 Lamm 11] , who competed with [12 12 Perot 12]
to head the ticket for [13 13 Perot 13] 's party , had told
them he would definitely not endorse [14 14 Perot 14] , but
they did not know whether he would endorse another
```

candidate . An announcement was planned in [15 15 Chicago 15] Wednesday . [16 16 Lamm 16] , 60 , is a three-term [17 17 Colorado 17] governor who left office in 1987 and vied for the [1 8 18 Reform Party 18] nomination after becoming disillusioned with both the [19 19 Democratic 19] and [20 20 Republican 20] parties . [21 21 Lamm 21] is a friend of President [22 22 Clinton 22] and supported him in the 1992 election . [23 23 Perot 23] won his party 's official nomination as its presidential candidate in a secret ballot earlier this month .

1: 949828 Richard_Lamm
2: 43505 Ross_Perot
3: 108956 Washington,_D.C.
4: 5399 Colorado
5: 5043544 Democratic_Party_(United_States)
6: 949828 Richard_Lamm
7: 43505 Ross_Perot
8: 295504 Reform_Party_of_the_United_States_of_America
9: 62028 CNN
10: 62028 CNN
11: 949828 Richard_Lamm
12: 43505 Ross_Perot
13: 43505 Ross_Perot
14: 43505 Ross_Perot
15: 6886 Chicago
16: 949828 Richard_Lamm
17: 5399 Colorado
18: 295504 Reform_Party_of_the_United_States_of_America
19: 5043544 Democratic_Party_(United_States)
20: 32070 Republican_Party_(United_States)
21: 949828 Richard_Lamm
22: 3356 Bill_Clinton
23: 43505 Ross_Perot
1: 949828 Richard_Lamm
 [pem_pos=0, final_scores=0.5916272,
final_local_scores=0.58675724, pem=0.0, similarity=21.59972]
2: 43505 Ross_Perot
 [pem_pos=0, final_scores=0.48420691,
final_local_scores=0.4416849, pem=0.0, similarity=17.270472]
3: 108956 Washington,_D.C.
 [pem_pos=0, final_scores=0.5633356,
final_local_scores=0.5926004, pem=-1.7496481,
similarity=31.288948]
4: 5399 Colorado
 [pem_pos=0, final_scores=0.6820795,
final_local_scores=0.75635254, pem=-0.26926568,
similarity=28.125092]
5: 5043544 Democratic_Party_(United_States)
 [pem_pos=0, final_scores=0.52508986,
final_local_scores=0.5084675, pem=-0.076093495,
similarity=19.677206]
6: 949828 Richard_Lamm
 [pem_pos=0, final_scores=0.49655622,

```
final_local_scores=0.47815806, pem=0.0, similarity=18.358904]
7: 43505 Ross_Perot
    [pem_pos=0, final_scores=0.4700641,
final_local_scores=0.42329752, pem=0.0, similarity=16.721756]
8: 295504 Reform_Party_of_the_United_States_of_America
    [pem_pos=0, final_scores=0.42962497,
final_local_scores=0.41778368, pem=-0.8355943,
similarity=21.101301]
9: 62028 CNN
    [pem_pos=0, final_scores=0.3204968,
final_local_scores=0.37027776, pem=0.0, similarity=15.13954]
10: 62028 CNN
    [pem_pos=0, final_scores=0.30897325,
final_local_scores=0.35769004, pem=0.0, similarity=14.763897]
11: 949828 Richard_Lamm
    [pem_pos=0, final_scores=0.63823044,
final_local_scores=0.6376642, pem=0.0, similarity=23.118885]
12: 43505 Ross_Perot
    [pem_pos=0, final_scores=0.5291366,
final_local_scores=0.49076378, pem=0.0, similarity=18.735083]
13: 43505 Ross_Perot
    [pem_pos=0, final_scores=0.48240632,
final_local_scores=0.439718, pem=0.0, similarity=17.211775]
14: 43505 Ross_Perot
    [pem_pos=0, final_scores=0.44313735,
final_local_scores=0.39682263, pem=0.0, similarity=15.931692]
15: 6886 Chicago
    [pem_pos=0, final_scores=0.58285725,
final_local_scores=0.63699275, pem=-0.20658566,
similarity=24.222292]
16: 949828 Richard_Lamm
    [pem_pos=0, final_scores=0.64735407,
final_local_scores=0.6476304, pem=0.0, similarity=23.416296]
17: 5399 Colorado
    [pem_pos=0, final_scores=0.47250825,
final_local_scores=0.5274278, pem=-0.26926568,
similarity=21.29352]
18: 295504 Reform_Party_of_the_United_States_of_America
    [pem_pos=0, final_scores=0.46984297,
final_local_scores=0.46171576, pem=-0.8355943,
similarity=22.412323]
19: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.45038193,
final_local_scores=0.42686045, pem=-0.076093495,
similarity=17.24189]
20: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.40562195,
final_local_scores=0.37212187, pem=-0.1310191,
similarity=15.907074]
21: 949828 Richard_Lamm
    [pem_pos=0, final_scores=0.6511784,
final_local_scores=0.6518079, pem=0.0, similarity=23.54096]
```

```
22: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.52642137,
final_local_scores=0.52257484, pem=-1.920079,
similarity=30.126078]
23: 43505 Ross_Perot
    [pem_pos=0, final_scores=0.5172128,
final_local_scores=0.47773886, pem=0.0, similarity=18.346394]
```

Sample: b'762_RESEARCH'

RESEARCH ALERT - Career Horizons said cut . -- [1 1 Donaldson Lufkin & Jenrette 1] cut its rating on Career Horizons Inc to market perform from outperform , according to market sources . -- Further details were not immediately available . -- The stock was up 3/4 at 35-7/8 .

```
1: 1520690 Donaldson,_Lufkin_&_Jenrette
1: 1520690 Donaldson,_Lufkin_&_Jenrette
    [pem_pos=0, final_scores=0.36938715,
final_local_scores=0.67608654, pem=0.0, similarity=24.265484]
```

Sample: b'763_Northern'

Northern States Power Co sets payout . [1 1 MINNEAPOLIS 1] 1996-08-28 Quarterly Latest Prior Amount \$ 0.69 \$ 0.69 Pay Oct 20 Record Oct 1 -- [2 2 Chicago 2] newsdesk 312 408-8787

```
1: 6097240 Minneapolis
2: 6886 Chicago
1: 6097240 Minneapolis
    [pem_pos=0, final_scores=0.50780576,
final_local_scores=0.45941907, pem=-0.0128788,
similarity=17.869732]
2: 6886 Chicago
    [pem_pos=0, final_scores=0.69454545,
final_local_scores=0.6634038, pem=-0.20658566,
similarity=25.010452]
```

Sample: b'764_US'

US investors mull appeal of Lloyd 's decision . Patricia Vowinkel [1 1 NEW YORK 1] 1996-08-27 [2 2 U.S. 2] investors in troubled Lloyd 's of London were considering late on Tuesday whether to appeal a [3 3 U.S. 3] court decision in favour of Lloyd 's and pledged to continue pursuing other legal actions . A [4 4 U.S. 4] appeals court on Tuesday gave Lloyd 's a reprieve , throwing out an injunction that the insurance giant said could have led to its collapse . A lower court issued the injunction on Friday and ordered Lloyd 's to give investors , known as Names , more information before requiring them to decide whether to accept a settlement offer as part of a reorganisation plan . " My prediction is that the Names will appeal , " said Kenneth Chiate , a [5 5 U.S. 5] Name and a chief negotiator for the American Names Association . " At this point , it is a sufficiently important decision that I 'm confident that they will appeal , " he said . " But to say that they definitely will

would be premature until we determine what the exact basis for the court 's ruling is . " Under the reorganisation plan , Lloyd 's plans to reinsure its massive liabilities into a new company called [6 6 Equitas 6] . The arrangement calls for investors to make additional payments to fund [7 7 E quitas 7] but also provides them with 3.2 billion stg in compensation to help reduce their prior outstanding liabilities . The Names had been scheduled to decide whether to accept or reject the offer by 1100 [8 8 GMT 8] Wednesday , but Lloyd 's chairman David Rowland said on Tuesday the offer would be extended . At least eight [9 9 U.S. 9] states have still some form of litigation pending , said John Head , spokesman for the Association of Lloyd 's State Chairmen , a group representing [10 10 U.S. 10] Names . " It goes without saying that we 're rather disappointed , " Head said of the decision by the [11 11 U.S. Court of Appeals 11] for the [12 1 2 Fourth Circuit 12] , sitting in [13 13 Baltimore 13] . But , he said , " we still have hope that somebody is going to see our point of view in this . " The [14 14 Colorado 14] attorney general told Lloyd 's last week it was considering a new legal action against the [15 15 British 15] insurance market , based on allegations of consumer fraud . " We have notified them of our concerns and asked them to give us a response , " said [16 16 Colorado 16] attorney general [17 17 Gale Norton 17] . [18 18 Norton 18] said she was concerned that the Lloyd 's agreement immunizes it from future litigation regarding [19 19 Equitas 19] and requires that all legal actions be heard outside of [20 20 Colorado 20] . In addition , she said she was concerned the plan may not offer investors enough protection from additional , future liabilities . Meanwhile , an appeal of a lawsuit filed by some 600 Names in [21 21 California 21] is still pending before the [22 22 U.S. Court of Appeals 22] for the Ninth Circuit , Chiate of the American Names Association said . That lawsuit , which seeks rescission and restitution , was dismissed in a [23 23 U.S. 23] district court . That case " is what I call the most significant alternate remedy for us , " Chiate said . The individual Names , however , now must decide whether to accept Lloyd 's settlement offer or reject the offer and pursue litigation . Chiate said he has advised Names that " they must make , individually , a risk benefit analysis . " Rejection involved forfeiting the compensation offer and risking the possibility of owing Lloyd 's two to three times more than Lloyd 's is now willing to accept , he said . But , in rejecting the offer , the Names would retain their rights to pursue litigation , he said .

- 1: 645042 New_York_City
- 2: 3434750 United_States
- 3: 3434750 United_States
- 4: 3434750 United_States
- 5: 3434750 United_States
- 6: 4659852 Equitas

```
7: 4659852 Equitas
8: 12701 Greenwich_Mean_Time
9: 3434750 United_States
10: 3434750 United_States
11: 70228 United_States_courts_of_appeals
12: 428324
United_States_Court_of_Appeals_for_the_Fourth_Circuit
13: 26997138 Baltimore
14: 5399 Colorado
15: 31717 United_Kingdom
16: 5399 Colorado
17: 285196 Gale_Norton
18: 285196 Gale_Norton
19: 4659852 Equitas
20: 5399 Colorado
21: 5407 California
22: 70228 United_States_courts_of_appeals
23: 3434750 United_States
1: 645042 New_York_City
    [pem_pos=1, final_scores=0.40141648,
final_local_scores=0.38513345, pem=-1.7906687,
similarity=25.320797]
2: 3434750 United_States
    [pem_pos=0, final_scores=0.43567365,
final_local_scores=0.3969285, pem=-0.14075686,
similarity=16.70031]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.48803908,
final_local_scores=0.45412976, pem=-0.14075686,
similarity=18.407307]
4: 3434750 United_States
    [pem_pos=0, final_scores=0.45866436,
final_local_scores=0.4220423, pem=-0.14075686,
similarity=17.449755]
5: 3434750 United_States
    [pem_pos=0, final_scores=0.57898825,
final_local_scores=0.55347794, pem=-0.14075686,
similarity=21.372057]
6: 4659852 Equitas
    [pem_pos=0, final_scores=0.6588389,
final_local_scores=0.81665844, pem=0.0, similarity=28.46043]
7: 4659852 Equitas
    [pem_pos=0, final_scores=0.5250641,
final_local_scores=0.6705298, pem=0.0, similarity=24.099659]
8: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.53151244,
final_local_scores=0.6555937, pem=0.0, similarity=23.653936]
9: 3434750 United_States
    [pem_pos=0, final_scores=0.50054204,
final_local_scores=0.46778733, pem=-0.14075686,
similarity=18.814878]
10: 3434750 United_States
```

```
[pem_pos=0, final_scores=0.5717135,
final_local_scores=0.54553133, pem=-0.14075686,
similarity=21.134914]
11: 70228 United_States_courts_of_appeals
    [pem_pos=1, final_scores=0.25788122,
final_local_scores=0.2673367, pem=-0.8717787,
similarity=16.808441]
12: 428324
United_Court_of_Appeals_for_the_Fourth_Circuit
    [pem_pos=0, final_scores=0.23104452,
final_local_scores=0.2600943, pem=0.0, similarity=11.851446]
13: 26997138 Baltimore
    [pem_pos=0, final_scores=0.56194,
final_local_scores=0.5725816, pem=-0.19949137,
similarity=22.261555]
14: 5399 Colorado
    [pem_pos=0, final_scores=0.5208543,
final_local_scores=0.52332133, pem=-0.26926568,
similarity=21.170975]
15: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.50276893,
final_local_scores=0.5058841, pem=-0.48936236,
similarity=21.847532]
16: 5399 Colorado
    [pem_pos=0, final_scores=0.58251715,
final_local_scores=0.5906787, pem=-0.26926568,
similarity=23.181053]
17: 285196 Gale_Norton
    [pem_pos=0, final_scores=0.68089575,
final_local_scores=0.75493735, pem=0.0, similarity=26.61855]
18: 285196 Gale_Norton
    [pem_pos=0, final_scores=1.2328397,
final_local_scores=1.3701875, pem=0.0, similarity=44.97884]
19: 4659852 Equitas
    [pem_pos=0, final_scores=0.59471405,
final_local_scores=0.7466118, pem=0.0, similarity=26.370098]
20: 5399 Colorado
    [pem_pos=0, final_scores=0.52698845,
final_local_scores=0.530022, pem=-0.26926568,
similarity=21.370937]
21: 5407 California
    [pem_pos=0, final_scores=0.5823633,
final_local_scores=0.58459145, pem=-0.08662026,
similarity=22.006144]
22: 70228 United_States_courts_of_appeals
    [pem_pos=1, final_scores=0.2879706,
final_local_scores=0.30020478, pem=-0.8717787,
similarity=17.789291]
23: 3434750 United_States
    [pem_pos=0, final_scores=0.5426917,
final_local_scores=0.5138294, pem=-0.14075686,
similarity=20.188864]
```

Sample: b'765_German'

[1 1 German 1] police made [2 2 Nazi 2] gestures , officials say . [3 3 NUREMBERG 3] , [4 4 Germany 4] 1996-08-28 [5 5 German 5] riot police made [6 6 Nazi 6] gestures at a private function earlier this month and may face dismissal for their actions , the [7 7 Bavarian 7] Interior Ministry said on Wednesday . The ministry declined to detail gestures the Nuremberg-based policemen had made at the August 13 function but added seven had been suspended from duty pending an internal inquiry . A spokesman for public prosecutors in the southern city , where dictator [8 8 Adolf Hitler 8] held some of his most infamous [1 1 Nazi 1] party rallies in the 1930s , said there were no plans to prosecute the officers as the gestures had not been made in public .

1: 11867 Germany
2: 21212 Nazi_Germany
3: 21287 Nuremberg
4: 11867 Germany
5: 11867 Germany
6: 21212 Nazi_Germany
7: 3764 Bavaria
8: 2731583 Adolf_Hitler
1: 21736 Nazi_Party

[winner: 21212 Nazi_Germany, pem_pos=1,
final_scores=0.07775481, final_local_scores=0.0005726218,
pem=-1.9176292, similarity=14.535167]
gt: 21736 Nazi_Party, pem_pos=2,
final_scores=-0.058056414, final_local_scores=-0.13110605,
pem=-2.3363276, similarity=12.88256]

1: 11867 Germany
[pem_pos=1, final_scores=0.8403931,
final_local_scores=0.8256369, pem=-1.0465823,
similarity=34.419846]

2: 21212 Nazi_Germany
[pem_pos=1, final_scores=0.24832629,
final_local_scores=0.18689606, pem=-1.9176292,
similarity=20.09543]

3: 21287 Nuremberg
[pem_pos=0, final_scores=0.5149907,
final_local_scores=0.4953345, pem=-0.14965801,
similarity=19.685345]

4: 11867 Germany
[pem_pos=0, final_scores=0.862886,
final_local_scores=0.85020703, pem=-0.15529114,
similarity=30.306086]

5: 11867 Germany
[pem_pos=1, final_scores=0.8778351,
final_local_scores=0.8665367, pem=-1.0465823,
similarity=35.640373]

6: 21212 Nazi_Germany
[pem_pos=1, final_scores=0.22208045,

```
final_local_scores=0.15822646, pem=-1.9176292,
similarity=19.239872]
7: 3764 Bavaria
    [pem_pos=0, final_scores=0.508927,
final_local_scores=0.48685694, pem=-0.05476965,
similarity=18.91634]
8: 2731583 Adolf_Hitler
    [pem_pos=0, final_scores=0.4742387,
final_local_scores=0.46211863, pem=0.0, similarity=17.880255]
1: 21212 Nazi_Germany
    [pem_pos=1, final_scores=0.07775481,
final_local_scores=0.0005726218, pem=-1.9176292,
similarity=14.535167]
```

Sample: b'766_Gun-wielding'

Gun-wielding motorist snapped by cool passenger . [1 1 BERLIN
1] 1996-08-28 A motorist threatened a fellow driver with a starting pistol as he overtook him illegally in the inside lane on a motorway near [2 2 Berlin 2] and was photographed in the act by the driver 's wife , prosecutors said on Wednesday . Prosecutors in the city of [3 3 Potsdam 3] said the 32-year-old man drew alongside the other car at about 110 kph (70 mph) and aimed his pistol at the driver . But the driver 's wife kept her nerve , got out her camera and photographed him . The man has been charged with dangerous driving , coercion and threatening behaviour .

```
1: 3354 Berlin
2: 3354 Berlin
3: 49112 Potsdam
1: 3354 Berlin
    [pem_pos=0, final_scores=0.6143703,
final_local_scores=0.55496866, pem=-0.15483057,
similarity=21.493078]
2: 3354 Berlin
    [pem_pos=0, final_scores=0.5487417,
final_local_scores=0.48327935, pem=-0.15483057,
similarity=19.353725]
3: 49112 Potsdam
    [pem_pos=0, final_scores=0.4419946,
final_local_scores=0.40182257, pem=-0.09432414,
similarity=16.59385]
```

Sample: b'767_German'

[1 1 German 1] prosecutors file sex tourism charges . [2 2 BE RLIN 2] 1996-08-28 [3 3 Berlin 3] prosecutors said on Wednesday they had filed charges against two [4 4 German 4] men for sexually abusing children in [5 5 Thailand 5] and distributing pornographic films and pictures of their degrading acts . The case is one of only a handful in which authorities have managed to track down suspects under a law which lets them pursue [6 6 Germans 6] who commit sex offences abroad . The pair , identified only as 43-year-old

clerk Dieter U and businessman Thomas S , 33 , are alleged to have carried out acts of sexual indecency with children as young as 10 years old between 1994 and 1995 . Their videos included pictures of one of the accused tying up a [7 7 Thai 7] boy and performing acts of sadistic torture on him , prosecutors said in a statement . In another scene , a young girl performed oral sex with an unidentified adult man . The new law was introduced with much fanfare in 1993 . But prosecutors face huge difficulties in gathering evidence and bringing witnesses to testify in a [8 8 German 8] court , and only one person has so far been convicted under the law . Investigators are probing several other cases . The [9 9 Berlin 9] prosecutors said they had been alerted to the two men by customs officials who intercepted packages containing pornographic photographs and order forms .

1: 11867 Germany
2: 3354 Berlin
3: 3354 Berlin
4: 11867 Germany
5: 30128 Thailand
6: 11867 Germany
7: 30128 Thailand
8: 11867 Germany
9: 3354 Berlin
1: 11867 Germany
[pem_pos=2, final_scores=0.94131666,
final_local_scores=0.91133505, pem=-1.0465823,
similarity=36.977245]
2: 3354 Berlin
[pem_pos=0, final_scores=0.59478194,
final_local_scores=0.55637777, pem=-0.15483057,
similarity=21.53513]
3: 3354 Berlin
[pem_pos=0, final_scores=0.54054207,
final_local_scores=0.4971289, pem=-0.15483057,
similarity=19.767025]
4: 11867 Germany
[pem_pos=2, final_scores=0.9419164,
final_local_scores=0.9119901, pem=-1.0465823,
similarity=36.996796]
5: 30128 Thailand
[pem_pos=0, final_scores=0.61544305,
final_local_scores=0.661575, pem=-0.24029273,
similarity=25.13918]
6: 11867 Germany
[pem_pos=1, final_scores=0.639777,
final_local_scores=0.57968014, pem=-1.5806074,
similarity=29.984116]
7: 30128 Thailand
[pem_pos=2, final_scores=0.6349043,
final_local_scores=0.6828335, pem=-1.7504287,
similarity=33.985928]

```
8: 11867 Germany
    [pem_pos=1, final_scores=0.88961786,
final_local_scores=0.85259324, pem=-1.0465823,
similarity=35.224274]
9: 3354 Berlin
    [pem_pos=0, final_scores=0.5910445,
final_local_scores=0.55229515, pem=-0.15483057,
similarity=21.413296]

Sample: b'768_France'
[1 1 France 1] 's Juppe on official visit to [2 2 Greece 2]
Sep 15 . [3 3 ATHENS 3] 1996-08-28 [4 4 French 4] Premier [5 5
Alain Juppe 5] will pay an official visit to [6 6 Greece 6]
on September 15 to celebrate 150 years of the French
Archaeological Society , government spokesman Dimitris Reppas
said on Wednesday . Juppe will meet [7 7 Greek 7] Prime
Minister [8 8 Costas Simitis 8] and Foreign Minister [9 9 The
odoros Pangalos 9] , Reppas told reporters . " The [10 10 Fre
nch 10] premier 's visit was planned to coincide with the
Archaeological Society 's celebrations . The [11 11 Greek 11]
government was asked for an official meeting with the prime
minister and the foreign minister and it said yes , " Reppas
said .
1: 5843419 France
2: 12108 Greece
3: 1216 Athens
4: 5843419 France
5: 434931 Alain_Juppé
6: 12108 Greece
7: 12108 Greece
8: 393824 Costas_Simitis
9: 2865375 Theodoros_Pangalos
10: 5843419 France
11: 12108 Greece
1: 5843419 France
    [pem_pos=0, final_scores=0.6641571,
final_local_scores=0.65436643, pem=-0.16240834,
similarity=24.500513]
2: 12108 Greece
    [pem_pos=0, final_scores=0.9067679,
final_local_scores=0.89154047, pem=-0.27607638,
similarity=32.196407]
3: 1216 Athens
    [pem_pos=0, final_scores=0.4516117,
final_local_scores=0.4417513, pem=-0.28831622,
similarity=18.840363]
4: 5843419 France
    [pem_pos=0, final_scores=0.71449894,
final_local_scores=0.70935726, pem=-0.80223274,
similarity=29.621016]
5: 434931 Alain_Juppé
    [pem_pos=0, final_scores=0.6952905,
```

```
final_local_scores=0.8176876, pem=0.0, similarity=28.49114]
6: 12108 Greece
    [pem_pos=0, final_scores=0.77421194,
final_local_scores=0.74674314, pem=-0.27607638,
similarity=27.875364]
7: 12108 Greece
    [pem_pos=1, final_scores=0.7686685,
final_local_scores=0.7406877, pem=-2.3247876,
similarity=38.83587]
8: 393824 Costas_Simitis
    [pem_pos=0, final_scores=0.6476673,
final_local_scores=0.74098396, pem=0.0, similarity=26.202152]
9: 2865375 Theodoros_Pangalos
    [pem_pos=0, final_scores=0.40444505,
final_local_scores=0.4796723, pem=0.0, similarity=18.40409]
10: 5843419 France
    [pem_pos=0, final_scores=0.69674414,
final_local_scores=0.6899628, pem=-0.80223274,
similarity=29.042246]
11: 12108 Greece
    [pem_pos=1, final_scores=0.846896,
final_local_scores=0.82613945, pem=-2.3247876,
similarity=41.385918]
```

Sample: b'769_Stork'

Stork H1 results breakdown per sector . [1 1 AMSTERDAM 1]
1996-08-28 First 24 weeks 1996 (millions of guilders unless otherwise stated) Industrial systems and components - Turnover 756 vs 829 - Operating profit 46 vs 48 - New orders received 876 vs 933 - Order book (billions) 1.07 vs 0.98 Industrial services - Turnover 657 vs 700 - Operating profit 9 vs 3 - New orders received (billions) 1.00 vs 1.09 - Order book (billions) 2.37 vs 2.01 NOTE - Order book figures refer to value of orders on books at end of period . -- [2 2 Amsterdam 2] newsroom +31 20 504 5000 , Fax +31 20 504 5040

1: 844 Amsterdam

2: 844 Amsterdam

1: 844 Amsterdam

```
    [pem_pos=0, final_scores=0.70334274,
final_local_scores=0.6402037, pem=-0.0642803,
similarity=23.544235]
```

2: 844 Amsterdam

```
    [pem_pos=0, final_scores=0.65845114,
final_local_scores=0.5911665, pem=-0.0642803,
similarity=22.080868]
```

Sample: b'770_Stephanie'

Stephanie of [1 1 Monaco 1] 's husband snapped cavorting . [2 2 ROME 2] 1996-08-28 Two [3 3 Italian 3] magazines published pictures on Wednesday of [4 4 Daniel Ducruet 4] , Princess Stephanie of [5 5 Monaco 5] 's husband and former bodyguard ,

cavorting naked with another woman by a poolside in [6 6 France 6] . The magazines , Eva Tremila and its sister publication [7 7 Gente 7] , printed up to 26 pages of photos of the woman undressing [8 8 Ducruet 8] , the pair embracing on a sunbed and finally both naked . Eva Tremila said other even more explicit photos were taken but it did not print them . The magazines named the woman as Fili Houteman , a 26-year-old [9 9 French 9] singer and dancer in a [10 10 Belgian 10] cabaret club . The photographs , an [11 11 Italian 11] exclusive , raised eyebrows in the tiny principality , where Stephanie 's father [12 12 Prince Rainier 12] , had long disapproved of his daughter 's choice of husband . Stephanie had two children with [13 13 Ducruet 13] before their marriage in July last year . Stephanie , [14 14 Caroline 14] and 1 Albert are the children of [15 15 Rainier 15] and former [16 16 Hollywood 16] screen goddess [17 17 Grace Kelly 17] , who was killed in a car crash in 1982 . " We have seen the photos but for the moment the palace has no comment , " a spokeswoman for [18 18 Prince Rainier 18] told [19 19 Reuters 19] . The magazines said the photographs were taken in Cap de Villefranche , some 15 km (nine miles) from [20 20 Monte Carlo 20] . [21 21 Gente 21] said [22 22 Ducruet 22] , a keen racing driver , met Houteman during a race in [23 23 Belgium 23] and photographers had been on their trail ever since . The magazine said video cameras had also been used to film the couple and that a sound-track existed .

1: 19261 Monaco
2: 25458 Rome
3: 14532 Italy
4: 1708370 Daniel_Ducruet
5: 19261 Monaco
6: 5843419 France
7: 19843098 Gente_(magazine)
8: 1708370 Daniel_Ducruet
9: 5843419 France
10: 3343 Belgium
11: 14532 Italy
12: 154987 Rainier_III,_Prince_of_Monaco
13: 1708370 Daniel_Ducruet
14: 292511 Caroline,_Princess_of_Hanover
15: 154987 Rainier_III,_Prince_of_Monaco
16: 53849 Hollywood
17: 42620 Grace_Kelly
18: 154987 Rainier_III,_Prince_of_Monaco
19: 18998750 Reuters
20: 137721 Monte_Carlo
21: 19843098 Gente_(magazine)
22: 1708370 Daniel_Ducruet
23: 3343 Belgium
1: 235952 Albert_II,_Prince_of_Monaco
[winner: 46755 George_VI, pem_pos=4,
final_scores=-0.42665306, final_local_scores=-0.46018207,

pem=-3.4064755, similarity=8.881906 |
RECALL MISS: 235952 Albert_II,_Prince_of_Monaco]
1: 19261 Monaco
[pem_pos=1, final_scores=0.4444192,
final_local_scores=0.425947, pem=-0.58243024,
similarity=19.968168]
2: 25458 Rome
[pem_pos=0, final_scores=0.48307008,
final_local_scores=0.4976291, pem=-0.2517334,
similarity=20.308922]
3: 14532 Italy
[pem_pos=1, final_scores=0.6953538,
final_local_scores=0.73041826, pem=-0.75948656,
similarity=30.017056]
4: 1708370 Daniel_Ducruet
[pem_pos=0, final_scores=0.67398787,
final_local_scores=0.7153827, pem=0.0, similarity=25.43816]
5: 19261 Monaco
[pem_pos=1, final_scores=0.4544487,
final_local_scores=0.4369027, pem=-0.58243024,
similarity=20.295107]
6: 5843419 France
[pem_pos=0, final_scores=0.70260644,
final_local_scores=0.71872824, pem=-0.16240834,
similarity=26.421198]
7: 19843098 Gente_(magazine)
[pem_pos=1, final_scores=0.5667114,
final_local_scores=0.71049166, pem=-1.3198198,
similarity=32.469585]
8: 1708370 Daniel_Ducruet
[pem_pos=0, final_scores=0.72764903,
final_local_scores=0.78395647, pem=0.0, similarity=27.484537]
9: 5843419 France
[pem_pos=0, final_scores=0.67072606,
final_local_scores=0.68390375, pem=-0.80223274,
similarity=28.861431]
10: 3343 Belgium
[pem_pos=0, final_scores=0.6746349,
final_local_scores=0.7031532, pem=-0.12403572,
similarity=25.747732]
11: 14532 Italy
[pem_pos=0, final_scores=0.57658845,
final_local_scores=0.60068506, pem=-0.75948656,
similarity=26.14556]
12: 154987 Rainier_III,_Prince_of_Monaco
[pem_pos=0, final_scores=0.5450716,
final_local_scores=0.5350928, pem=0.0, similarity=20.05795]
13: 1708370 Daniel_Ducruet
[pem_pos=0, final_scores=0.76235896,
final_local_scores=0.82187176, pem=0.0, similarity=28.616005]
14: 292511 Caroline,_Princess_of_Hanover
[pem_pos=11, final_scores=0.6221722,

```
final_local_scores=0.6536057, pem=-3.6799214,
similarity=43.606594]
15: 154987 Rainier_III,_Prince_of_Monaco
    [pem_pos=0, final_scores=0.96852034,
final_local_scores=1.004385, pem=0.0, similarity=34.06256]
16: 53849 Hollywood
    [pem_pos=1, final_scores=0.37578088,
final_local_scores=0.40833312, pem=-2.0106025,
similarity=27.209158]
17: 42620 Grace_Kelly
    [pem_pos=0, final_scores=0.55561507,
final_local_scores=0.59198976, pem=0.0, similarity=21.755869]
18: 154987 Rainier_III,_Prince_of_Monaco
    [pem_pos=0, final_scores=0.5045492,
final_local_scores=0.49082822, pem=0.0, similarity=18.737005]
19: 18998750 Reuters
    [pem_pos=0, final_scores=0.3390832,
final_local_scores=0.40814358, pem=0.0, similarity=16.269531]
20: 137721 Monte_Carlo
    [pem_pos=0, final_scores=0.43026316,
final_local_scores=0.43758804, pem=-0.81665266,
similarity=21.589298]
21: 19843098 Gente_(magazine)
    [pem_pos=1, final_scores=0.46544534,
final_local_scores=0.59987384, pem=-1.3198198,
similarity=29.16853]
22: 1708370 Daniel_Ducruet
    [pem_pos=0, final_scores=0.6994194,
final_local_scores=0.7531199, pem=0.0, similarity=26.564312]
23: 3343 Belgium
    [pem_pos=0, final_scores=0.62969,
final_local_scores=0.6540578, pem=-0.24180195,
similarity=24.92306]
```

Sample: b'771_Highlights'

Highlights of Wednesday 's [1 1 Commission 1] briefing . [2 2 BRUSSELS 2] 1996-08-28 Following are highlights of the midday briefing by the [3 3 European Commission 3] on Wednesday : In response to a question , [4 4 Commission 4] spokesman Joao Vale de Almeida said there had been no developments regarding the [5 5 Commission 5] 's position concerning the dispute with [6 6 Germany 6] and [7 7 Saxony 7] over state aid to [8 8 Volkswagen 8] . He said there was some possibility of further talks with [9 9 Germany 9] before the next [10 10 Commission 1 0] meeting of September 4 . - - - - The [11 11 Commission 11] released the following documents : - IP / 96 / 804 : [12 12 Commission 12] approves acquisition of Pao de Acucar by [13 13 Auchan 13] . - IP / 96 / 805 : [14 14 Commission 14] finds acquisition of CAMAT by AGF-IART does not fall under the merger regulation . - IP / 96 / 806 : [15 15 Commission 15] clears acquisition of [16 16 Austrian 16] food retail chain [1 7 17 Billa 17] by [18 18 German 18] group Rewe-

Handelsgruppe . - SPEECH / 96 / 202 : Speech by [19 19 Europe
an 19] Commissioner [20 20 Anita Grdin 20] at the World
Congress against Sexual Exploitation of Children in [21 21 St
ockholm 21] . - [22 22 Eurostat 22] news release 51/96 :
March-May 1996 EU industrial production figures .

1: 9974 European_Commission
2: 3708 Brussels
3: 9974 European_Commission
4: 9974 European_Commission
5: 9974 European_Commission
6: 11867 Germany
7: 28395 Saxony
8: 32413 Volkswagen
9: 11867 Germany
10: 9974 European_Commission
11: 9974 European_Commission
12: 9974 European_Commission
13: 1981235 Auchan
14: 9974 European_Commission
15: 9974 European_Commission
16: 26964606 Austria
17: 307855 Billa_(supermarket)
18: 11867 Germany
19: 9317 European_Union
20: 2346687 Anita_Gradin
21: 26741 Stockholm
22: 679519 Eurostat
1: 9974 European_Commission
[pem_pos=0, final_scores=0.67021793,
final_local_scores=0.64104164, pem=-0.76247066,
similarity=27.366108]
2: 3708 Brussels
[pem_pos=0, final_scores=0.55547446,
final_local_scores=0.5529748, pem=-0.025440974,
similarity=20.729937]
3: 9974 European_Commission
[pem_pos=0, final_scores=0.5323698,
final_local_scores=0.49046355, pem=0.0, similarity=18.726124]
4: 9974 European_Commission
[pem_pos=0, final_scores=0.5698584,
final_local_scores=0.5314141, pem=-0.76247066,
similarity=24.094604]
5: 9974 European_Commission
[pem_pos=0, final_scores=0.61462975,
final_local_scores=0.58032, pem=-0.76247066,
similarity=25.554054]
6: 11867 Germany
[pem_pos=0, final_scores=0.9337478,
final_local_scores=0.97205514, pem=-0.15529114,
similarity=33.942276]
7: 28395 Saxony
[pem_pos=0, final_scores=0.47419626,

```
final_local_scores=0.53237677, pem=-0.14660512,
similarity=20.77416]
8: 32413 Volkswagen
    [pem_pos=0, final_scores=0.2572493,
final_local_scores=0.29427052, pem=-0.0033464932,
similarity=12.889531]
9: 11867 Germany
    [pem_pos=0, final_scores=0.6851095,
final_local_scores=0.7004556, pem=-0.15529114,
similarity=25.837204]
10: 9974 European_Commission
    [pem_pos=0, final_scores=0.66228175,
final_local_scores=0.6323726, pem=-0.76247066,
similarity=27.107409]
11: 9974 European_Commission
    [pem_pos=0, final_scores=0.6777175,
final_local_scores=0.6492338, pem=-0.76247066,
similarity=27.61058]
12: 9974 European_Commission
    [pem_pos=0, final_scores=0.773608,
final_local_scores=0.7539796, pem=-0.76247066,
similarity=30.736403]
13: 1981235 Auchan
    [pem_pos=0, final_scores=0.41508454,
final_local_scores=0.52421206, pem=0.0, similarity=19.733246]
14: 9974 European_Commission
    [pem_pos=0, final_scores=0.6893595,
final_local_scores=0.6619509, pem=-0.76247066,
similarity=27.990082]
15: 9974 European_Commission
    [pem_pos=0, final_scores=0.6667295,
final_local_scores=0.6372311, pem=-0.76247066,
similarity=27.252396]
16: 26964606 Austria
    [pem_pos=0, final_scores=0.5944511,
final_local_scores=0.604629, pem=-0.40674555,
similarity=24.344994]
17: 307855 Billa_(supermarket)
    [pem_pos=1, final_scores=0.76252323,
final_local_scores=0.94851464, pem=-0.8482281,
similarity=37.00808]
18: 11867 Germany
    [pem_pos=1, final_scores=0.6971465,
final_local_scores=0.71360415, pem=-1.0465823,
similarity=31.076563]
19: 9317 European_Union
    [pem_pos=1, final_scores=0.26271105,
final_local_scores=0.2073381, pem=-2.9373715,
similarity=26.250978]
20: 2346687 Anita_Gradin
    [pem_pos=0, final_scores=0.5007039,
final_local_scores=0.6600235, pem=0.0, similarity=23.78613]
```

```
21: 26741 Stockholm
    [pem_pos=0, final_scores=0.6590673,
final_local_scores=0.70024914, pem=-0.1047564,
similarity=25.556225]
22: 679519 Eurostat
    [pem_pos=0, final_scores=0.41505522,
final_local_scores=0.46687078, pem=0.0, similarity=18.022068]
```

Sample: b'772_Spanish'

[1 1 Spanish 1] tomato warriors paint the town red . [2 2 BU
NOL 2] , [3 3 Spain 3] 1996-08-28 Revellers painted the town
red on Wednesday as the 1996 edition of the world 's biggest
tomato fight began in the eastern [4 4 Spanish 4] village of [
5 5 Bunol 5] . Thousands of people pelted each other with
armfuls of ripe tomatoes as streets , walls and windows were
coated in a blood-red wash . A single firework after midday
signalled the start of the fruit-throwing frenzy , during
which participants hurl some 100 tonnes of tomatoes trucked
in for the occasion . Local historians say the tradition
began in 1945 when disgruntled locals began spontaneously to
bombard the priest and mayor at the annual fiesta in [6 6 Bun
ol 6] (pronounced Boo-nee-OL) . The festival 's fame has
grown and now attracts between 15,000 and 20,000 people ,
many of them foreigners .

```
1: 26667 Spain
2: 277229 Buñol
3: 26667 Spain
4: 26667 Spain
5: 277229 Buñol
6: 277229 Buñol
1: 26667 Spain
    [pem_pos=1, final_scores=0.5818538,
final_local_scores=0.5886675, pem=-1.2425895,
similarity=28.41412]
2: 277229 Buñol
    [pem_pos=0, final_scores=0.39375985,
final_local_scores=0.3832034, pem=0.0, similarity=15.525266]
3: 26667 Spain
    [pem_pos=0, final_scores=0.5885131,
final_local_scores=0.5959417, pem=-0.18857051,
similarity=22.899279]
4: 26667 Spain
    [pem_pos=1, final_scores=0.5392894,
final_local_scores=0.5421723, pem=-1.2425895,
similarity=27.026611]
5: 277229 Buñol
    [pem_pos=0, final_scores=0.4866485,
final_local_scores=0.4846701, pem=0.0, similarity=18.553236]
6: 277229 Buñol
    [pem_pos=0, final_scores=0.46319908,
final_local_scores=0.4590552, pem=0.0, similarity=17.788836]
```

Sample: b'773_PRESS'

PRESS DIGEST - [1 1 GREECE 1] - AUG 28 . [2 2 ATHENS 2]
1996-08-28 Leading stories in the [3 3 Greek 3] financial
press : IMERISIA -- Pre-election debate heats up on economic
issues as conservative [4 4 New Democracy 4] party promises
seven measures includind tax relief for farmers and socialist
[5 5 Pasok 5] defends progress on economic convergence with
the EU -- Finance ministry scrambles to find temporary
solution to regulation which slaps a 15 percent tax rate on
gains from trading of bonds and coupons by mutual funds --
Finance ministry will cut 12-month T-bill rate by 10 basis
points to 12.70 percent in the upcoming end August issue
FINANCIAL KATHIMERINI -- Inflows of more than \$ 500 million
are seen in the interbank market and the bourse in the last
three days reflecting confidence in the post-election
economic policy -- [6 6 Athens Metro 6] subway project hits
snags which could delay delivery to the year 2000 and
overshoot the original budgeted cost of 520 billion drachmas
-- State National Bank of Greece will start real auction
programme September 9 to lighten up on its real estate
holdings KERDOS -- [7 7 New Democracy 7] leader [8 8 Miltiadis
Evert 8] vows support mesures for farmers and small business
as he kicks off the conservative party 's campaign --
National Economy Minister Yannos Papandoniou defends " hard
drachma " foreign exchange policy , says it wo n't change
EXPRESS -- Message of unity from the conservative [9 9 New De
mocracy 9] party as former prime minister [10 10 Constantine M
itsotakis 10] and [11 11 Miltiadis Evert 11] shake hands
NAFTEMBORIKI -- Government defends " hard drachma " policy ,
says it will continue unchanged after the elections --
Conservative opposition [12 12 New Democracy 12] promises
series of measures on the economy 30 days after the elections
aiming at 4.0 percent GDP growth rate annually -- George
Georgopoulos , Athens Newsroom +301 3311812-4

1: 12108 Greece

2: 1216 Athens

3: 12108 Greece

4: 424225 New_Democracy_(Greece)

5: 405559 Panhellenic_Socialist_Movement

6: 430783 Athens_Metro

7: 424225 New_Democracy_(Greece)

8: 5481658 Miltiadis_Evert

9: 424225 New_Democracy_(Greece)

10: 464282 Konstantinos_Mitsotakis

11: 5481658 Miltiadis_Evert

12: 424225 New_Democracy_(Greece)

1: 12108 Greece

[pem_pos=0, final_scores=0.534349,
final_local_scores=0.5462912, pem=-0.27607638,
similarity=21.893476]

2: 1216 Athens

[pem_pos=0, final_scores=0.4522184,

```
final_local_scores=0.47545767, pem=-0.28831622,
similarity=19.846228]
3: 12108 Greece
    [pem_pos=1, final_scores=0.7006568,
final_local_scores=0.7279571, pem=-2.3247876,
similarity=38.455963]
4: 424225 New_Democracy_(Greece)
    [pem_pos=0, final_scores=0.4827749,
final_local_scores=0.4351679, pem=-0.35315335,
similarity=18.996494]
5: 405559 Panhellenic_Socialist_Movement
    [pem_pos=0, final_scores=0.47941965,
final_local_scores=0.4561779, pem=0.0, similarity=17.70297]
6: 430783 Athens_Metro
    [pem_pos=0, final_scores=0.42267293,
final_local_scores=0.5308625, pem=0.0, similarity=19.93171]
7: 424225 New_Democracy_(Greece)
    [pem_pos=0, final_scores=0.42228204,
final_local_scores=0.36908865, pem=-0.35315335,
similarity=17.024557]
8: 5481658 Miltiadis_Evert
    [pem_pos=0, final_scores=0.6379785,
final_local_scores=0.7154665, pem=0.0, similarity=25.44066]
9: 424225 New_Democracy_(Greece)
    [pem_pos=0, final_scores=0.4939807,
final_local_scores=0.44740856, pem=-0.35315335,
similarity=19.36178]
10: 464282 Konstantinos_Mitsotakis
    [pem_pos=0, final_scores=0.5742169,
final_local_scores=0.58214825, pem=0.0, similarity=21.462177]
11: 5481658 Miltiadis_Evert
    [pem_pos=0, final_scores=0.52309996,
final_local_scores=0.5899791, pem=0.0, similarity=21.695866]
12: 424225 New_Democracy_(Greece)
    [pem_pos=0, final_scores=0.44417065,
final_local_scores=0.3929987, pem=-0.35315335,
similarity=17.738081]
```

Sample: b'774_HOEK'

HOEK LOOS H1 NET PROFIT 28.9 MLN GUILDER^S . [1 1 AMSTERDAM 1]
1996-08-28 First half 1996 (in millions of guilders unless otherwise stated) Net per shr (guilders) 4.38 vs 3.70 Net profit 28.9 vs 24.5 Turnover 273.6 vs 290.3 Operating profit 44.4 vs 40.7 Note - Industrial gases maker Hoek Loos NV . Interest charges 2.20 vs 5.05 Tax 13.26 vs 11.16 -- [2 2 Amsterdam 2] newsroom +31 20 504 5000 , Fax +31 20 504 5040

1: 844 Amsterdam

2: 844 Amsterdam

1: 844 Amsterdam

```
[pem_pos=0, final_scores=0.73104686,
final_local_scores=0.6704663, pem=-0.0642803,
similarity=24.447332]
```

2: 844 Amsterdam
[pem_pos=0, final_scores=0.6712449,
final_local_scores=0.6051417, pem=-0.0642803,
similarity=22.497915]

Sample: b'775_Sstagecoach'
[1 1 Stagecoach 1] sees [2 2 Swebus 2] deal agreed next
week . [3 3 LONDON 3] 1996-08-28 [4 4 British 4] bus and
passenger rail operator Stagecoach Holdings Plc said on
Wednesday that its negotiations to acquire [5 5 Swedish 5]
long distance bus operator Swebus AB were set to lead to a
signed agreement next week . Four weeks ago [6 6 Stagecoach 6]
said it had agreed the deal in principle , and it expected to
pay 110 million stg-plus for the firm , with [7 7 Swebus 7]
current owner , the state railway company . " The directors
report that negotiations with the vendors of Swebus AB are
proceeding and they expect an agreement (conditional on
shareholder approval) will be signed next week , " [8 8 Stag
ecoach 8] said in a statement . -- London Newsroom +44 171
542 7717

1: 1040592 Stagecoach_Group
2: 232036 Nobina_Sverige
3: 17867 London
4: 31717 United_Kingdom
5: 5058739 Sweden
6: 1040592 Stagecoach_Group
7: 232036 Nobina_Sverige
8: 1040592 Stagecoach_Group
1: 1040592 Stagecoach_Group
[pem_pos=1, final_scores=0.42916983,
final_local_scores=0.40954256, pem=-1.4079248,
similarity=23.967794]

2: 232036 Nobina_Sverige
[pem_pos=0, final_scores=0.6236121,
final_local_scores=0.70613796, pem=0.0, similarity=25.162277]

3: 17867 London
[pem_pos=0, final_scores=0.642948,
final_local_scores=0.7059501, pem=-0.09086737,
similarity=25.650822]

4: 31717 United_Kingdom
[pem_pos=0, final_scores=0.45736313,
final_local_scores=0.5061171, pem=-0.48936236,
similarity=21.854486]

5: 5058739 Sweden
[pem_pos=0, final_scores=0.5493656,
final_local_scores=0.63609636, pem=-0.63135606,
similarity=26.50551]

6: 1040592 Stagecoach_Group
[pem_pos=1, final_scores=0.3757422,
final_local_scores=0.35118103, pem=-1.4079248,
similarity=22.22617]

7: 232036 Nobina_Sverige

```
[pem_pos=0, final_scores=0.57790107,
final_local_scores=0.6562056, pem=0.0, similarity=23.672197]
8: 1040592 Stagecoach_Group
[pem_pos=1, final_scores=0.28435224,
final_local_scores=0.25135136, pem=-1.4079248,
similarity=19.247053]
```

Sample: b'776_NZ'

NZ bills gain ground after see-saw session . [1 1 WELLINGTON
1] 1996-08-28 0515 [2 2 GMT 2] The [3 3 New Zealand 3] money market gained slightly at Wednesday 's close after what dealers described as a see-saw trading session . Ninety-day bank bill rates shed five points to 9.93 percent and September bank bill futures rose four to 90.18 . However , bonds finished largely flat . " Our bonds were better bid initially but they sold off on a lack of demand , and the short end went with it too on a lower currency . " There were big buyers at the base of where the market sold to , and when the currency got bought back on talk of a samurai the market got bought back again , " a dealer said . Volumes were reasonable in the money market but thin in bonds . Dealers said the market seemed to be trading a range and would wait for more political polls to provide direction . They were confident of further eurokiwi issuance but said the timing was less of a certainty . -- [4 4 Wellington 4] newsroom (64
4) 473 4746

1: 33804 Wellington

2: 12701 Greenwich_Mean_Time

3: 4913064 New_Zealand

4: 33804 Wellington

1: 33804 Wellington

```
[pem_pos=0, final_scores=0.5023632,
final_local_scores=0.46704912, pem=-0.9265605,
similarity=23.06617]
```

2: 12701 Greenwich_Mean_Time

```
[pem_pos=0, final_scores=0.62839067,
final_local_scores=0.7701932, pem=0.0, similarity=27.073814]
```

3: 4913064 New_Zealand

```
[pem_pos=0, final_scores=0.5773017,
final_local_scores=0.5809833, pem=-0.23795462,
similarity=22.721447]
```

4: 33804 Wellington

```
[pem_pos=0, final_scores=0.5088096,
final_local_scores=0.4740908, pem=-0.9265605,
similarity=23.276308]
```

Sample: b'777_RTRS'

RTRS - Guinness Peat expects strong full yr . [1 1 SYDNEY 1]
1996-08-28 [2 2 British-based 2] investment company Guinness Peat Group Plc (GPG) said on Wednesday it expected a strong full year result . " We think we 're in a position to produce a strong result , however a lot of our profitability must

inevitably depend on a number of (company) results , " said GPG director Garry Weiss . GPG earlier said its net profit for the six months to June 30 rose to 9.77 million pounds from 6.93 million in the previous first half . The company did not declare an interim dividend as in the previous year . Weiss said the [3 3 Australian 3] share market had been somewhat negative for much of 1996 and this had some effect on the company 's first half results . " However , it certainly is a very pleasing result for the first six months , " he said . Weiss said most of the company 's half year earning stemmed from the sale of its 50 percent stake in Physicians Insurance Co of [4 4 Ohio 4] . He said the company decided to sell its [5 5 U.S. 5] investment in order to consolidate investments closer to its administrative base . GPG said its stakes in Tyndall Australia Ltd and Mid-East Minerals Ltd both contributed strongly to GPG 's first half earnings . -- Sydney Newsroom 61-2 9373-1800

1: 27862 Sydney
2: 31717 United_Kingdom
3: 4689264 Australia
4: 22199 Ohio
5: 3434750 United_States
1: 27862 Sydney
[pem_pos=0, final_scores=0.59572124,
final_local_scores=0.59698206, pem=-0.1443489,
similarity=22.68984]
2: 31717 United_Kingdom
[pem_pos=0, final_scores=0.4782452,
final_local_scores=0.44998294, pem=0.0, similarity=17.5181]
3: 4689264 Australia
[pem_pos=0, final_scores=0.71176964,
final_local_scores=0.7009976, pem=-0.491009,
similarity=27.679062]
4: 22199 Ohio
[pem_pos=0, final_scores=0.7107379,
final_local_scores=0.73795813, pem=-0.19147073,
similarity=27.153103]
5: 3434750 United_States
[pem_pos=0, final_scores=0.47255325,
final_local_scores=0.427635, pem=-0.14075686,
similarity=17.616653]

Sample: b'778_RTRS'

RTRS - Newcrest Q4 net profit A\$ 4.3 mln . [1 1 SYDNEY 1]
1996-08-28 Gold miner [2 2 Newcrest Mining Ltd 2] said on Wednesday it posted a A\$ 4.3 million profit after tax in the final quarter of the year to June 30 , 1996 . Earlier , Newcrest reported a drop in net profit after abnormalities to A\$ 20.81 million for the year from A\$ 42.4 million the previous year . Newcrest said earnings from the [3 3 Telfer 3] and [4 4 Boddington 4] mines were lower than the previous year due to lower head grades at the mines , forcing gold production

lower . Production costs also rose eight percent during the year to A\$ 406 per ounce .

1: 27862 Sydney
2: 1798576 Newcrest_Mining
3: 22088081 Telfer_Mine
4: 22097576 Boddington_Gold_Mine
1: 27862 Sydney
[pem_pos=0, final_scores=0.5012392,
final_local_scores=0.61262476, pem=-0.1443489,
similarity=23.15665]
2: 1798576 Newcrest_Mining
[pem_pos=0, final_scores=0.40979475,
final_local_scores=0.42188454, pem=0.0, similarity=16.679588]
3: 22088081 Telfer_Mine
[pem_pos=5, final_scores=0.42235106,
final_local_scores=0.4591155, pem=-2.6383092,
similarity=32.138172]
4: 22097576 Boddington_Gold_Mine
[pem_pos=4, final_scores=0.41669095,
final_local_scores=0.49490708, pem=-2.5742197,
similarity=32.857735]

Sample: b'779_RTRS'

RTRS - [1 1 Queensland 1] gunman evades police in bush hunt . [2 2 BRISBANE 2] 1996-08-28 [3 3 Australian 3] police on Wednesday continued to hunt a gunman in dense bushland after he killed his wife and wounded three other people , warning the man is extremely dangerous and may take a hostage to escape . The shooting occurred around 6.30 a.m. (2030 [4 4 GMT 4]) on Tuesday at Glenwood , south of Maryborough , about 150 km (93 miles) north of [5 5 Brisbane 5] on the [6 6 Queensland 6] state coast . Police have declared an " emergent situation " in the area , giving them powers to raid houses , search cars , close schools , quarantine the area and evacuate people . " It is one step short of an emergency situation , " a police spokesman said via telephone from a command post in the bush . " We have not had any sightings , but we suspect he is armed , possibly with a .22 rifle and / or a self-loading shotgun . He is considered extremely dangerous , " he said . " It 's a possibility , not a probability , he may take a hostage , but we have measures in place if that is the case . " William Fox broke into his wife 's home on Tuesday morning , shooting her dead and wounding his 16-year-old son , his son 's girlfriend and a neighbour , police said . All three wounded are in a satisfactory condition in hospital . Fox initially fled from the shooting in a car , but then abandoned the car and entered dense bushland . Fox is a skilled bushman who knows the area very well , police said . About 60 police , helicopters and fixed-wing aircraft have maintained an overnight cordon around 15 sq km (six sq miles) of bush near Glenwood . The area is littered with caves and police believed Fox has a hideout

which has enabled him to evade capture . [7 7 Australia 7] 's six states and two territories are involved in heated debate over the introduction of tough new firearm laws , including the banning rapid fire weapons , after a shooting massacre in the island state of [8 8 Tasmania 8] . On April 28 , a lone gunman went on a shooting rampage at the site of the historic [9 9 Port Arthur 9] penal settlement , killing 35 people .

1: 59051 Queensland
2: 192093 Brisbane
3: 4689264 Australia
4: 12701 Greenwich_Mean_Time
5: 192093 Brisbane
6: 59051 Queensland
7: 4689264 Australia
8: 29944 Tasmania
9: 265537 Port_Arthur,_Tasmania
1: 59051 Queensland
 [pem_pos=0, final_scores=0.5344216,
final_local_scores=0.4901191, pem=-0.083265886,
similarity=19.168657]
2: 192093 Brisbane
 [pem_pos=0, final_scores=0.5206435,
final_local_scores=0.4891705, pem=-0.15081817,
similarity=19.507708]
3: 4689264 Australia
 [pem_pos=0, final_scores=0.7414411,
final_local_scores=0.7123942, pem=-0.491009,
similarity=28.01916]
4: 12701 Greenwich_Mean_Time
 [pem_pos=0, final_scores=0.5811191,
final_local_scores=0.71744275, pem=0.0, similarity=25.499636]
5: 192093 Brisbane
 [pem_pos=0, final_scores=0.44716746,
final_local_scores=0.40890908, pem=-0.15081817,
similarity=17.112547]
6: 59051 Queensland
 [pem_pos=0, final_scores=0.54208577,
final_local_scores=0.498491, pem=-0.083265886,
similarity=19.418491]
7: 4689264 Australia
 [pem_pos=0, final_scores=0.61200434,
final_local_scores=0.5710042, pem=-0.17168477,
similarity=22.063267]
8: 29944 Tasmania
 [pem_pos=0, final_scores=0.58510226,
final_local_scores=0.5522068, pem=-0.048622325,
similarity=20.833082]
9: 265537 Port_Arthur,_Tasmania
 [pem_pos=2, final_scores=0.36873156,
final_local_scores=0.43894798, pem=-1.4185004,
similarity=24.902822]

Sample: b'780_Shanghai'

Shanghai Ek Chor opens new motorcycle plant . [1 1 SHANGHAI 1] 1996-08-28 Shanghai-Ek Chor Motorcycle Co , a Sino-Thai joint venture , opened a new plant to produce gasoline engines in the [2 2 Pudong New Area 2] of [3 3 Shanghai 3] , the [4 4 Xinhua 4] news agency reported on Wednesday . The plant , requiring an investment of three billion baht , has a floor space of 50,000 square metres and is designed to produce 600,000 gasoline engines a year , to be sold in [5 5 China 5] , [6 6 South America 6] , the [7 7 Middle East 7] and [8 8 Africa 8] , it said . Capacity is expected to reach one million engines by the year 2000 , it said . Shanghai-Ek Chor is jointly owned by the Shanghai Automobile Corporation and Ek Chor China Motorcycle . It started operations in January 1985 and has registered capital of 1.56 billion baht , it said but gave no further details . The joint venture has two motorcycle plants making Xingfu motorcycles and aims to be [9 9 China 9] 's biggest producer by the year 2000 , with output of two million units .

1: 27643 Shanghai
2: 402737 Pudong
3: 27643 Shanghai
4: 263168 Xinhua_News_Agency
5: 5405 China
6: 26769 South_America
7: 19323 Middle_East
8: 5334607 Africa
9: 5405 China
1: 27643 Shanghai
[pem_pos=0, final_scores=0.756687,
final_local_scores=0.75470126, pem=-0.10072064,
similarity=27.159239]
2: 402737 Pudong
[pem_pos=0, final_scores=0.44083476,
final_local_scores=0.4990058, pem=0.0, similarity=18.981041]
3: 27643 Shanghai
[pem_pos=0, final_scores=0.60120374,
final_local_scores=0.58485943, pem=-0.10072064,
similarity=22.090818]
4: 263168 Xinhua_News_Agency
[pem_pos=0, final_scores=0.48769963,
final_local_scores=0.53617287, pem=0.0, similarity=20.090181]
5: 5405 China
[pem_pos=0, final_scores=0.52016693,
final_local_scores=0.47017723, pem=-0.15669827,
similarity=18.972889]
6: 26769 South_America
[pem_pos=0, final_scores=0.45629042,
final_local_scores=0.47021008, pem=0.0, similarity=18.12172]
7: 19323 Middle_East
[pem_pos=0, final_scores=0.4720623,
final_local_scores=0.47634393, pem=0.0, similarity=18.304766]

```
8: 5334607 Africa
    [pem_pos=0, final_scores=0.5636328,
final_local_scores=0.56979746, pem=-0.13197383,
similarity=21.8113]
9: 5405 China
    [pem_pos=0, final_scores=0.48658723,
final_local_scores=0.43349648, pem=-0.15669827,
similarity=17.878262]

Sample: b'781_Khmer'
Khmer Rouge 's [1 1 Ieng Sary 1] confirms break with [2 2 Pol
Pot 2] . ARANYAPRATHET , [3 3 Thailand 3] 1996-08-28
Dissident Khmer Rouge leader [4 4 Ieng Sary 4] confirmed on
Wednesday that he had broken with [5 5 Pol Pot 5] and other
hardliners of the guerrilla group and had formed a rival
movement . [6 6 Ieng Sary 6] said in a written statement ,
the first since his split with [7 7 Pol Pot 7] earlier this
month , that the new movement to be called the Democratic
National United Movement ( DNUM ) would seek an end to civil
war and work towards reconciliation with the [8 8 Cambodian 8]
government . " I would like to inform you about my decision
to break away from [9 9 Pol Pot 9] , [10 10 Ta Mok 10] , [11 1
1 Son Sen 11] 's dictatorial group , " he said in a copy of
the statement obtained by [12 12 Reuters 12] . " We believe
that our country will be reduced to nothing if the [13 13 Khm
er 13] people continue to fight against each other
indefinitely .... For this reason we decided to break away
from that dictatorial group and found a movement named ' '
Democratic National United Movement ' , " he said . [14 14 Ie
ng Sary 14] was sentenced to death in absentia for his role
in the mass genocide in [15 15 Cambodia 15] during the Khmer
Rouge rule of terror between 1975-1979 when over a million
people were executed or died of starvation , disease or
overwork in mass labour camps . The French-educated , former
brother-in-law of [16 16 Pol Pot 16] was foreign minister in
the Khmer Rouge government that ruled [17 17 Cambodia 17] and
was seen as the group 's second in command .
1: 1160532 Ieng_Sary
2: 24326 Pol_Pot
3: 30128 Thailand
4: 1160532 Ieng_Sary
5: 24326 Pol_Pot
6: 1160532 Ieng_Sary
7: 24326 Pol_Pot
8: 334751 Cambodia
9: 24326 Pol_Pot
10: 222456 Ta_Mok
11: 868969 Son_Sen
12: 18998750 Reuters
13: 1429661 Khmer_people
14: 1160532 Ieng_Sary
15: 334751 Cambodia
```

```
16: 24326 Pol_Pot
17: 334751 Cambodia
1: 1160532 Ieng_Sary
    [pem_pos=0, final_scores=0.62640554,
final_local_scores=0.62309724, pem=0.0, similarity=22.684177]
2: 24326 Pol_Pot
    [pem_pos=0, final_scores=0.68781114,
final_local_scores=0.6480097, pem=0.0, similarity=23.427616]
3: 30128 Thailand
    [pem_pos=0, final_scores=0.594215,
final_local_scores=0.66122276, pem=-0.24029273,
similarity=25.128668]
4: 1160532 Ieng_Sary
    [pem_pos=0, final_scores=0.6278694,
final_local_scores=0.6246964, pem=0.0, similarity=22.7319]
5: 24326 Pol_Pot
    [pem_pos=0, final_scores=0.7652922,
final_local_scores=0.73559755, pem=0.0, similarity=26.04141]
6: 1160532 Ieng_Sary
    [pem_pos=0, final_scores=0.71710473,
final_local_scores=0.7221723, pem=0.0, similarity=25.640776]
7: 24326 Pol_Pot
    [pem_pos=0, final_scores=0.71049386,
final_local_scores=0.67573863, pem=0.0, similarity=24.255102]
8: 334751 Cambodia
    [pem_pos=0, final_scores=0.49650556,
final_local_scores=0.5010773, pem=-0.33141,
similarity=20.845118]
9: 24326 Pol_Pot
    [pem_pos=0, final_scores=0.72333086,
final_local_scores=0.6897611, pem=0.0, similarity=24.67356]
10: 222456 Ta_Mok
    [pem_pos=0, final_scores=0.57273006,
final_local_scores=0.5908137, pem=0.0, similarity=21.720772]
11: 868969 Son_Sen
    [pem_pos=0, final_scores=0.63145846,
final_local_scores=0.64726484, pem=0.0, similarity=23.405386]
12: 18998750 Reuters
    [pem_pos=0, final_scores=0.36747867,
final_local_scores=0.45331562, pem=0.0, similarity=17.617556]
13: 1429661 Khmer_people
    [pem_pos=1, final_scores=0.19413324,
final_local_scores=0.19582212, pem=-1.1484975,
similarity=16.179144]
14: 1160532 Ieng_Sary
    [pem_pos=0, final_scores=0.65881926,
final_local_scores=0.65850437, pem=0.0, similarity=23.740797]
15: 334751 Cambodia
    [pem_pos=0, final_scores=0.5338052,
final_local_scores=0.5418215, pem=-0.27299485,
similarity=21.743336]
16: 24326 Pol_Pot
```

```
[pem_pos=0, final_scores=0.7591775,
final_local_scores=0.7289181, pem=0.0, similarity=25.842083]
17: 334751 Cambodia
    [pem_pos=0, final_scores=0.4014842,
final_local_scores=0.39728093, pem=-0.27299485,
similarity=17.429956]
```

Sample: b'782_Two'

Two dead in [1 1 Cambodia 1] helicopter crash . [2 2 PHNOM PE NH 2] 1996-08-28 Two people were killed and six were injured after a helicopter crashed in bad weather in northern [3 3 Cambodia 3] , a government minister said on Wednesday . The 15 survivors who had been on board the Russian-made [4 4 MI-17 4] helicopter were taken to hospital from the remote jungle crash site about 150 km (90 miles) north of [5 5 Phnom Penh 5] , Information Minister [6 6 Ieng Mouly 6] said . The cause of the crash of the helicopter , which went down on Sunday while on a routine resupply flight between [7 7 Phnom Penh 7] and [8 8 Stung Treng 8] , was not known . [9 9 Ieng Mouly 9] said the aircraft went down during a rain storm .

```
1: 334751 Cambodia
2: 51591 Phnom_Penh
3: 334751 Cambodia
4: 1493374 Mil_Mi-17
5: 51591 Phnom_Penh
6: 12219062 Ieng_Mouly
7: 51591 Phnom_Penh
8: 6658821 Stung_Treng
9: 12219062 Ieng_Mouly
1: 334751 Cambodia
    [pem_pos=0, final_scores=0.51668674,
final_local_scores=0.4929784, pem=-0.27299485,
similarity=20.28576]
2: 51591 Phnom_Penh
    [pem_pos=0, final_scores=0.483177,
final_local_scores=0.43517745, pem=0.0, similarity=17.076277]
3: 334751 Cambodia
    [pem_pos=0, final_scores=0.45471478,
final_local_scores=0.42528343, pem=-0.27299485,
similarity=18.265608]
4: 1493374 Mil_Mi-17
    [pem_pos=0, final_scores=0.3998835,
final_local_scores=0.51221687, pem=0.0, similarity=19.375286]
5: 51591 Phnom_Penh
    [pem_pos=0, final_scores=0.5455507,
final_local_scores=0.5033113, pem=0.0, similarity=19.109526]
6: 12219062 Ieng_Mouly
    [pem_pos=0, final_scores=0.51305467,
final_local_scores=0.63968027, pem=0.0, similarity=23.179049]
7: 51591 Phnom_Penh
    [pem_pos=0, final_scores=0.55911034,
final_local_scores=0.51812315, pem=0.0, similarity=19.55154]
```

```
8: 6658821 Stung_Treng
    [pem_pos=0, final_scores=0.26000726,
final_local_scores=0.3074826, pem=0.0, similarity=13.265607]
9: 12219062 Ieng_Mouly
    [pem_pos=0, final_scores=0.4724071,
final_local_scores=0.5952789, pem=0.0, similarity=21.854023]

Sample: b'783_MOF'
[1 1 MOF 1] 's Kubo says believes [2 2 BOJ 2] rate policy
unchanged . [3 3 TOKYO 3] 1996-08-28 [4 4 Japan 4] 's Finance
Minister Wataru Kubo told a news conference on Wednesday that
he believes that the [5 5 Bank of Japan 5] 's ( [6 6 BOJ 6] )
interest rate policy which is geared towards ensuring
economic growth has not changed after the release of the
central bank 's " tankan " survey . The [7 7 BOJ 7] released
the August tankan , its quarterly short-term corporate
survey , in the morning and it showed business outlook had
worsened . However , Kubo said it did not necessarily show a
substantial worsening of the economy . " The question is what
the [8 8 BOJ 8] is going to do with its interest rate
policy . The [9 9 BOJ 9] governor has made it clear that the [
10 10 BOJ 10] 's policy is aimed at ensuring basis for
economic recovery . I believe this policy has not changed , "
he said . Asked if a supplementary budget for 1996/97 was
needed to support the economy , Kubo said the tankan results
would not lead to any immediate decision on the need for an
extra budget . " I do n't think we should immediately draw a
conclusion that the economic recovery has come to a halt or
that signs of a economic contraction have emerged , " Kubo
said . " The economy is not recovering smoothly or at a fast
pace . " Kubo said he would make a decision on the need for a
supplementary budget after an announcement in mid-September
of [11 11 Japan 11] 's gross domestic product for the April-
June quarter . " I would like to see how the economy moved in
the first half of 1996 , " he said .
1: 870389 Ministry_of_Finance_(Japan)
2: 188155 Bank_of_Japan
3: 30057 Tokyo
4: 15573 Japan
5: 188155 Bank_of_Japan
6: 188155 Bank_of_Japan
7: 188155 Bank_of_Japan
8: 188155 Bank_of_Japan
9: 188155 Bank_of_Japan
10: 188155 Bank_of_Japan
11: 15573 Japan
1: 870389 Ministry_of_Finance_(Japan)
    [pem_pos=3, final_scores=0.38695818,
final_local_scores=0.39961076, pem=-1.609438,
similarity=24.76727]
2: 188155 Bank_of_Japan
    [pem_pos=0, final_scores=0.6405168,
```

```
final_local_scores=0.5926393, pem=0.0, similarity=21.775253]
3: 30057 Tokyo
    [pem_pos=0, final_scores=0.5924062,
final_local_scores=0.6614617, pem=-0.09283106,
similarity=24.33388]
4: 15573 Japan
    [pem_pos=0, final_scores=0.6338871,
final_local_scores=0.707409, pem=-0.13549899,
similarity=25.937073]
5: 188155 Bank_of_Japan
    [pem_pos=0, final_scores=0.7902331,
final_local_scores=0.7561816, pem=0.0, similarity=26.65568]
6: 188155 Bank_of_Japan
    [pem_pos=0, final_scores=0.5916414,
final_local_scores=0.5392504, pem=0.0, similarity=20.18202]
7: 188155 Bank_of_Japan
    [pem_pos=0, final_scores=0.73559767,
final_local_scores=0.69650066, pem=0.0, similarity=24.874681]
8: 188155 Bank_of_Japan
    [pem_pos=0, final_scores=0.7221486,
final_local_scores=0.68180966, pem=0.0, similarity=24.436274]
9: 188155 Bank_of_Japan
    [pem_pos=0, final_scores=0.63674515,
final_local_scores=0.58851933, pem=0.0, similarity=21.652304]
10: 188155 Bank_of_Japan
    [pem_pos=0, final_scores=0.622413,
final_local_scores=0.5728636, pem=0.0, similarity=21.185104]
11: 15573 Japan
    [pem_pos=0, final_scores=0.6027101,
final_local_scores=0.67335296, pem=-0.13549899,
similarity=24.920773]
```

Sample: b'784_China'

```
[1 1 China 1] says militant [2 2 Japan 2] must face war
past . [3 3 BEIJING 3] 1996-08-28 [4 4 China 4] on Wednesday
called on [5 5 Japan 5] to acknowledge its wartime past and
put a stop to a tide of resurgent militarism to prevent
similar atrocities in future . " Some [6 6 Japanese 6] are
still unrepentant about the atrocities committed by the [7 7 J
apanese 7] militarists during the war , " said a commentary
in the official [8 8 China Daily 8] . " If they are still
undecided whether the war [9 9 Japan 9] launched was
aggressive in nature , it will be difficult to tell whether
they will do the same again , " the newspaper said . [10 10 Ch
ina 10] raised indignant protests after several [11 11 Japane
se 11] cabinet ministers visited a shrine dedicated to their
country 's war dead on August 15 , the 51st anniversary of [12
12 Japan 12] 's [13 13 World War Two 13] surrender . "
Numerous [14 14 Japanese 14] politicians have tried to
whitewash their country 's war atrocities in recent years ,
the commentary said . [15 15 China 15] estimates 35 million [1
6 16 Chinese 16] were killed or wounded by invading [17 17 Ja
```

panese 17] troops from 1931 to 1945 . " The [18 18 Japanese 18] have never genuinely apologised for their wartime crimes , " the commentary said . [19 19 Japanese 19] Prime Minister [20 20 Ryutaro Hashimoto 20] marked the August 15 anniversary by expressing " remorse " for foreign victims of [21 21 Japan 21] 's [22 22 World War Two 22] atrocities .

1: 5405 China
2: 15573 Japan
3: 18603746 Beijing
4: 5405 China
5: 15573 Japan
6: 15573 Japan
7: 15573 Japan
8: 737850 China_Daily
9: 15573 Japan
10: 5405 China
11: 15573 Japan
12: 15573 Japan
13: 32927 World_War_II
14: 15573 Japan
15: 5405 China
16: 5405 China
17: 15573 Japan
18: 15573 Japan
19: 15573 Japan
20: 368549 Ryutaro_Hashimoto
21: 15573 Japan
22: 32927 World_War_II
1: 5405 China
 [pem_pos=0, final_scores=0.5510506,
final_local_scores=0.5191096, pem=-0.15669827,
similarity=20.433128]
2: 15573 Japan
 [pem_pos=0, final_scores=0.6538051,
final_local_scores=0.5930554, pem=-0.13549899,
similarity=22.524534]
3: 18603746 Beijing
 [pem_pos=0, final_scores=0.50123954,
final_local_scores=0.51703846, pem=-0.042161036,
similarity=19.748451]
4: 5405 China
 [pem_pos=0, final_scores=0.5517582,
final_local_scores=0.51988256, pem=-0.15669827,
similarity=20.456196]
5: 15573 Japan
 [pem_pos=0, final_scores=0.62696546,
final_local_scores=0.56373715, pem=-0.13549899,
similarity=21.649618]
6: 15573 Japan
 [pem_pos=0, final_scores=0.7573287,
final_local_scores=0.7067364, pem=-0.99076414,
similarity=30.568064]

7: 15573 Japan
[pem_pos=0, final_scores=0.8120678,
final_local_scores=0.76593333, pem=-0.99076414,
similarity=32.33462]
8: 737850 China_Daily
[pem_pos=0, final_scores=0.5062613,
final_local_scores=0.5894936, pem=0.0, similarity=21.681377]
9: 15573 Japan
[pem_pos=0, final_scores=0.59154993,
final_local_scores=0.52505106, pem=-0.13549899,
similarity=20.495148]
10: 5405 China
[pem_pos=0, final_scores=0.6043008,
final_local_scores=0.57727736, pem=-0.15669827,
similarity=22.16897]
11: 15573 Japan
[pem_pos=0, final_scores=0.8046096,
final_local_scores=0.7583837, pem=-0.99076414,
similarity=32.109325]
12: 15573 Japan
[pem_pos=0, final_scores=0.6321977,
final_local_scores=0.5694525, pem=-0.13549899,
similarity=21.820175]
13: 32927 World_War_II
[pem_pos=0, final_scores=0.4608261,
final_local_scores=0.49162006, pem=0.0, similarity=18.760635]
14: 15573 Japan
[pem_pos=0, final_scores=0.86264795,
final_local_scores=0.8217817, pem=-0.99076414,
similarity=34.001247]
15: 5405 China
[pem_pos=0, final_scores=0.48100036,
final_local_scores=0.44259036, pem=-0.15669827,
similarity=18.149641]
16: 5405 China
[pem_pos=1, final_scores=0.58791095,
final_local_scores=0.5593739, pem=-1.4196125,
similarity=28.50262]
17: 15573 Japan
[pem_pos=0, final_scores=0.8198115,
final_local_scores=0.7749894, pem=-0.99076414,
similarity=32.604874]
18: 15573 Japan
[pem_pos=0, final_scores=0.8121992,
final_local_scores=0.766674, pem=-0.99076414,
similarity=32.356724]
19: 15573 Japan
[pem_pos=0, final_scores=0.9695094,
final_local_scores=0.93851155, pem=-0.99076414,
similarity=37.4847]
20: 368549 Ryutaro_Hashimoto
[pem_pos=0, final_scores=0.72511643,

```
final_local_scores=0.86210424, pem=0.0, similarity=29.816624]
21: 15573 Japan
    [pem_pos=0, final_scores=0.81618696,
final_local_scores=0.77043295, pem=-0.13549899,
similarity=27.817833]
22: 32927 World_War_II
    [pem_pos=0, final_scores=0.511257,
final_local_scores=0.54670817, pem=0.0, similarity=20.404575]
```

Sample: b'785_Japan'

[1 1 Japan 1] coalition party leader plans to resign . [2 2 T
OKYO 2] 1996-08-28 The leader of a junior partner in [3 3 Jap
an 3] 's three-party ruling coalition plans to resign to
quell a political rebellion , party officials said on
Wednesday . The officials said [4 4 New Party Sakigake 4]
President [5 5 Masayoshi Takemura 5] , finance minister until
the beginning of this year , promised his resignation in a
meeting with the politician who set off the rebellion in the
smallest coalition member . They said the date of [6 6 Takemu
ra 6] 's resignation would be determined by party officials .
The [7 7 Sakigake 7] row has caused jitters in its coalition
partners , Prime Minister [8 8 Ryutaro Hashimoto 8] 's [9 9 Li
beral Democratic Party 9] ([10 10 LDP 10]) , the biggest
grouping , and the [11 11 Social Democratic Party 11] . But
analysts said the row was not expected to immediately
destabilise the government as even if [12 12 Sakigake 12]
splits apart it has so few seats a loss of support would not
lead to a general election . The dispute pits [13 13 Takemura
13] , who founded [14 14 Sakigake 14] in 1993 as a reform-
oriented [15 15 LDP 15] splinter group , against party
official [16 16 Yukio Hatoyama 16] , who says he will leave [1
7 17 Sakigake 17] to form a reformist political group next
month . [18 18 Hatoyama 18] , the 49-year-old grandson of a
1950s prime minister , on Tuesday quit as [19 19 Sakigake 19]
secretary general and has publicly snubbed the 62-year-old [20
20 Takemura 20] , pointedly ruling his mentor out as a
possible member of the new political force . Marathon talks
between the two former allies on Tuesday night and Wednesday
morning failed to resolve the dispute over the role of [21 21
Takemura 21] , seen by [22 22 Hatoyama 22] backers as tainted
by his senior role in the LDP-dominated coalition . The
presence of [23 23 Takemura 23] , whose role as finance
minister in passing an unpopular plan to use taxpayer funds
to wind up failed housing loan firms ruined his reputation as
a reformer , has stalled [24 24 Hatoyama 24] 's efforts to
attract to his new group defectors from the opposition camp ,
analysts said . Media reports say that at most 10 of 23 [25 25
Sakigake 25] members , joined by a handful of Social
Democrats , will follow [26 26 Hatoyama 26] when he bolts --
far short of the 50 lawmakers needed to topple [27 27 Hashimo
to 27] 's eight-month-old government . [28 28 Hashimoto 28]
-- who returns from a 10-day [29 29 Latin American 29] tour

on Saturday -- must call polls by mid-1997 , and has repeatedly said he would not call an early general election . But many analysts and politicians believe he may dissolve parliament soon after it reconvenes in early October .

1: 15573 Japan
2: 30057 Tokyo
3: 15573 Japan
4: 387050 New_Party_Sakigake
5: 23634266 Masayoshi_Takemura
6: 23634266 Masayoshi_Takemura
7: 387050 New_Party_Sakigake
8: 368549 Ryutaro_Hashimoto
9: 164231 Liberal_Democratic_Party_(Japan)
10: 164231 Liberal_Democratic_Party_(Japan)
11: 375454 Social_Democratic_Party_(Japan)
12: 387050 New_Party_Sakigake
13: 23634266 Masayoshi_Takemura
14: 387050 New_Party_Sakigake
15: 164231 Liberal_Democratic_Party_(Japan)
16: 2687269 Yukio_Hatoyama
17: 387050 New_Party_Sakigake
18: 2687269 Yukio_Hatoyama
19: 387050 New_Party_Sakigake
20: 23634266 Masayoshi_Takemura
21: 23634266 Masayoshi_Takemura
22: 2687269 Yukio_Hatoyama
23: 23634266 Masayoshi_Takemura
24: 2687269 Yukio_Hatoyama
25: 387050 New_Party_Sakigake
26: 2687269 Yukio_Hatoyama
27: 368549 Ryutaro_Hashimoto
28: 368549 Ryutaro_Hashimoto
29: 18524 Latin_America
1: 15573 Japan
[pem_pos=0, final_scores=0.4970305,
final_local_scores=0.57693046, pem=-0.13549899,
similarity=22.043333]
2: 30057 Tokyo
[pem_pos=0, final_scores=0.6126485,
final_local_scores=0.69824135, pem=-0.09283106,
similarity=25.431458]
3: 15573 Japan
[pem_pos=0, final_scores=0.58607644,
final_local_scores=0.6741996, pem=-0.13549899,
similarity=24.946037]
4: 387050 New_Party_Sakigake
[pem_pos=0, final_scores=0.5305216,
final_local_scores=0.54070747, pem=0.0, similarity=20.225502]
5: 23634266 Masayoshi_Takemura
[pem_pos=0, final_scores=0.6892831,
final_local_scores=0.71105486, pem=0.0, similarity=25.309008]
6: 23634266 Masayoshi_Takemura

```
[pem_pos=0, final_scores=0.89389974,
final_local_scores=0.93968564, pem=0.0, similarity=32.131805]
7: 387050 New_Party_Sakigake
    [pem_pos=1, final_scores=0.5800865,
final_local_scores=0.5948496, pem=-0.68746537,
similarity=25.579754]
8: 368549 Ryutaro_Hashimoto
    [pem_pos=0, final_scores=0.8592456,
final_local_scores=0.88420457, pem=0.0, similarity=30.476141]
9: 164231 Liberal_Democratic_Party_(Japan)
    [pem_pos=0, final_scores=0.53667915,
final_local_scores=0.50979644, pem=-0.41367084,
similarity=21.552662]
10: 164231 Liberal_Democratic_Party_(Japan)
    [pem_pos=0, final_scores=0.47056812,
final_local_scores=0.43758017, pem=-0.76990485,
similarity=21.33484]
11: 375454 Social_Democratic_Party_(Japan)
    [pem_pos=7, final_scores=0.28610963,
final_local_scores=0.26639056, pem=-3.0487192,
similarity=28.618748]
12: 387050 New_Party_Sakigake
    [pem_pos=1, final_scores=0.641125,
final_local_scores=0.6615249, pem=-0.68746537,
similarity=27.569479]
13: 23634266 Masayoshi_Takemura
    [pem_pos=0, final_scores=0.9380625,
final_local_scores=0.9879268, pem=0.0, similarity=33.57142]
14: 387050 New_Party_Sakigake
    [pem_pos=1, final_scores=0.53082967,
final_local_scores=0.541044, pem=-0.68746537,
similarity=23.974089]
15: 164231 Liberal_Democratic_Party_(Japan)
    [pem_pos=0, final_scores=0.4553538,
final_local_scores=0.42096084, pem=-0.76990485,
similarity=20.838888]
16: 2687269 Yukio_Hatoyama
    [pem_pos=0, final_scores=0.6313339,
final_local_scores=0.61710423, pem=0.0, similarity=22.505335]
17: 387050 New_Party_Sakigake
    [pem_pos=1, final_scores=0.5375164,
final_local_scores=0.54834825, pem=-0.68746537,
similarity=24.192062]
18: 2687269 Yukio_Hatoyama
    [pem_pos=0, final_scores=0.58421534,
final_local_scores=0.5690962, pem=0.0, similarity=21.07268]
19: 387050 New_Party_Sakigake
    [pem_pos=1, final_scores=0.64732784,
final_local_scores=0.6683006, pem=-0.68746537,
similarity=27.77168]
20: 23634266 Masayoshi_Takemura
    [pem_pos=0, final_scores=0.90896446,
```

```
final_local_scores=0.95614165, pem=0.0, similarity=32.622887]
21: 23634266 Masayoshi_Takemura
    [pem_pos=0, final_scores=0.9090716,
final_local_scores=0.9562587, pem=0.0, similarity=32.62638]
22: 2687269 Yukio_Hatoyama
    [pem_pos=0, final_scores=0.54010373,
final_local_scores=0.520911, pem=0.0, similarity=19.634735]
23: 23634266 Masayoshi_Takemura
    [pem_pos=0, final_scores=0.9589259,
final_local_scores=1.0107169, pem=0.0, similarity=34.25152]
24: 2687269 Yukio_Hatoyama
    [pem_pos=0, final_scores=0.564805,
final_local_scores=0.54789335, pem=0.0, similarity=20.439943]
25: 387050 New_Party_Sakigake
    [pem_pos=1, final_scores=0.7357791,
final_local_scores=0.7649201, pem=-0.68746537,
similarity=30.654999]
26: 2687269 Yukio_Hatoyama
    [pem_pos=0, final_scores=0.5600043,
final_local_scores=0.54264927, pem=0.0, similarity=20.28345]
27: 368549 Ryutaro_Hashimoto
    [pem_pos=0, final_scores=0.8900438,
final_local_scores=0.92231065, pem=0.0, similarity=31.613302]
28: 368549 Ryutaro_Hashimoto
    [pem_pos=0, final_scores=0.9967721,
final_local_scores=1.0388951, pem=0.0, similarity=35.092415]
29: 18524 Latin_America
    [pem_pos=0, final_scores=0.3203599,
final_local_scores=0.44213855, pem=-0.13736542,
similarity=18.031025]
```

Sample: b'786_Liu'

Liu Chong Hing interim net up 2.7 pct . [1 1 HONG KONG 1]
1996-08-28 Six months ended June 30 (in million [2 2 HK\$ 2]
unless stated) Shr ([3 3 H.K. 3] cents) 65.61 vs 63.87
Dividend ([4 4 H.K. 4] cents) 18.0 vs 18.0 Exceptional
items nil vs nil Net 249.53 vs 242.94 Turnover 119.49 vs
134.40 Company name Liu Chong Hing Investment Ltd Books close
September 23-27 Dividend payable October 8 NOTE - Liu Chong
Hing engages in property development and investment ,
warehousing , banking and insurance services . -- Hong Kong
Newsroom (852) 2843 6368

```
1: 13404 Hong_Kong
2: 309734 Hong_Kong_dollar
3: 13404 Hong_Kong
4: 13404 Hong_Kong
1: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.55698884,
final_local_scores=0.4935972, pem=-0.22402541,
similarity=20.037924]
2: 309734 Hong_Kong_dollar
    [pem_pos=0, final_scores=0.5545674,
```

```
final_local_scores=0.59737015, pem=0.0, similarity=21.91643]
3: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.6813058,
final_local_scores=0.62939465, pem=-0.6931472,
similarity=26.641546]
4: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.63950515,
final_local_scores=0.58373374, pem=-0.6931472,
similarity=25.278934]
```

Sample: b'787_Fire'

Fire bomb hurled at [1 1 U.S. 1] consulate in [2 2 Indonesia 2] . [3 3 JAKARTA 3] 1996-08-28 A fire bomb was thrown over the fence into the grounds of the [4 4 U.S. 4] Consulate-General in [5 5 Indonesia 5] 's second largest city of [6 6 Surabaya 6] but no one was hurt , a mission official said on Wednesday . Craig Stromme , [7 7 U.S. 7] embassy spokesman in [8 8 Jakarta 8] , 700 km (430 miles) west of [9 9 Surabaya 9] , confirmed the Tuesday morning attack . " Somebody threw a molotov cocktail over the fence and it went into the parking lot . It did n't hit anybody or anything , " Stromme said . He said there was no immediate explanation for the attack or any information on those responsible .

```
1: 3434750 United_States
2: 14579 Indonesia
3: 16275 Jakarta
4: 3434750 United_States
5: 14579 Indonesia
6: 212635 Surabaya
7: 3434750 United_States
8: 16275 Jakarta
9: 212635 Surabaya
1: 3434750 United_States
    [pem_pos=0, final_scores=0.52476496,
final_local_scores=0.5221794, pem=-0.14075686,
similarity=20.438046]
2: 14579 Indonesia
    [pem_pos=0, final_scores=0.54665834,
final_local_scores=0.5144102, pem=-0.23151328,
similarity=20.699745]
3: 16275 Jakarta
    [pem_pos=0, final_scores=0.51906306,
final_local_scores=0.47493392, pem=0.0, similarity=18.262688]
4: 3434750 United_States
    [pem_pos=0, final_scores=0.48721617,
final_local_scores=0.48116308, pem=-0.14075686,
similarity=19.214035]
5: 14579 Indonesia
    [pem_pos=0, final_scores=0.5943169,
final_local_scores=0.5664699, pem=-0.23151328,
similarity=22.253311]
6: 212635 Surabaya
```

```
[pem_pos=0, final_scores=0.44387025,
final_local_scores=0.42808068, pem=0.0, similarity=16.864494]
7: 3434750 United_States
    [pem_pos=0, final_scores=0.5594252,
final_local_scores=0.5600404, pem=-0.14075686,
similarity=21.567892]
8: 16275 Jakarta
    [pem_pos=0, final_scores=0.46197587,
final_local_scores=0.4125749, pem=0.0, similarity=16.40177]
9: 212635 Surabaya
    [pem_pos=0, final_scores=0.40723556,
final_local_scores=0.38806283, pem=0.0, similarity=15.670281]
```

Sample: b'788_Shanghai'

[1 1 Shanghai 1] novelist murdered at home . [2 2 SHANGHAI 2] 1996-08-28 A [3 3 Shanghai 3] novelist was murdered at her home on Sunday , an official of the city 's Writers Association said on Wednesday . The victim was Dai Houying , who wrote about [4 4 China 4] 's 1966-76 leftist [1 Cultural Revolution 1] and the lives of [5 5 Chinese 5] intellectuals , the official said . The killing was under investigation , she said . She gave no further details . Born in 1937 in the central province of [6 6 Anhui 6] , Dai came to [7 7 Shanghai 7] as a student and remained in the city as a prolific author and teacher of [8 8 Chinese 8] . She was divorced and lived alone , leaving one daughter who received university education in [9 9 Hawaii 9] and lives in [10 10 Chicago 10] , a friend said . Dai 's most famous book , " Ren A Ren " (People , People) , was translated into [11 11 German 11] and [12 12 English 12] , he said .

```
1: 27643 Shanghai
2: 27643 Shanghai
3: 27643 Shanghai
4: 5405 China
5: 5405 China
6: 105059 Anhui
7: 27643 Shanghai
8: 5405 China
9: 13270 Hawaii
10: 6886 Chicago
11: 11884 German_language
12: 8569916 English_language
1: 27643 Shanghai
    [pem_pos=0, final_scores=0.44481045,
final_local_scores=0.39001566, pem=-0.10072064,
similarity=16.276293]
2: 27643 Shanghai
    [pem_pos=0, final_scores=0.73542774,
final_local_scores=0.7074709, pem=-0.10072064,
similarity=25.74979]
3: 27643 Shanghai
    [pem_pos=0, final_scores=0.51286346,
```

```
final_local_scores=0.4643532, pem=-0.10072064,
similarity=18.494673]
4: 5405 China
    [pem_pos=0, final_scores=0.50537986,
final_local_scores=0.46692872, pem=-0.15669827,
similarity=18.875948]
5: 5405 China
    [pem_pos=1, final_scores=0.6893439,
final_local_scores=0.66788155, pem=-1.4196125,
similarity=31.740704]
6: 105059 Anhui
    [pem_pos=0, final_scores=0.36914045,
final_local_scores=0.38428992, pem=0.0, similarity=15.557691]
7: 27643 Shanghai
    [pem_pos=0, final_scores=0.5541404,
final_local_scores=0.509442, pem=-0.10072064,
similarity=19.840212]
8: 5405 China
    [pem_pos=1, final_scores=0.43682998,
final_local_scores=0.39204842, pem=-1.4196125,
similarity=23.509295]
9: 13270 Hawaii
    [pem_pos=0, final_scores=0.4030984,
final_local_scores=0.42200756, pem=-0.2312327,
similarity=17.94074]
10: 6886 Chicago
    [pem_pos=0, final_scores=0.6260985,
final_local_scores=0.66676325, pem=-0.20658566,
similarity=25.110704]
11: 11884 German_language
    [pem_pos=0, final_scores=0.34546566,
final_local_scores=0.39254093, pem=-0.981596,
similarity=21.141989]
12: 8569916 English_language
    [pem_pos=0, final_scores=0.5380711,
final_local_scores=0.5820847, pem=-0.640266,
similarity=24.942148]
1: 45311 Cultural_Revolution
    [pem_pos=0, final_scores=0.04891543,
final_local_scores=0.04273607, pem=-0.008948008,
similarity=5.413703]
```

Sample: b'789_Hwa'

Hwa Kay plunges on rights issue , earnings . [1 1 HONG KONG 1] 1996-08-28 Shares of Hwa Kay Thai Holdings Ltd plunged to an all-time low after the company announced a rights issue plan and also reported a sharp fall in earnings , brokers said . The stock fell [2 2 HK\$ 2] 0.23 , or 30.26 percent , to an all-time low of [3 3 HK\$ 3] 0.53 . " Investors unloaded their shares due to the poor earnings outlook following a sharp profit decline . The rights issue also prompted dilution fears , " said a dealing director at a local

```
brokerage .
1: 13404 Hong_Kong
2: 309734 Hong_Kong_dollar
3: 309734 Hong_Kong_dollar
1: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5183509,
final_local_scores=0.5578092, pem=-0.22402541,
similarity=21.954136]
2: 309734 Hong_Kong_dollar
    [pem_pos=0, final_scores=0.580519,
final_local_scores=0.53716695, pem=0.0, similarity=20.119846]
3: 309734 Hong_Kong_dollar
    [pem_pos=0, final_scores=0.6296408,
final_local_scores=0.590825, pem=0.0, similarity=21.72111]
```

Sample: b'790_Japan'

```
[1 1 Japan 1] July refined zinc imports off 47.5 pct yr / yr . [2 2 TOKYO 2] 1996-08-28 [3 3 Japan 3] 's refined zinc imports in July totalled 3,684 tonnes , off 47.5 percent from 7,011 tonnes in the same month a year earlier , according to [4 4 Ministry of Finance 4] data released on Wednesday . Figures were as follows ( in tonnes ) : July 96 June 96 July 95 Total 3,684 3,292 7,011 Major suppliers : [5 5 China 5] 961 1,683 5,539 Refined zinc imports in the first seven months of 1996 totalled 115,941 tonnes , up 38.4 percent from 83,801 tonnes in the year-ago period . -- Tokyo Commodities Desk ( 813 3432 6179 )
```

```
1: 15573 Japan
2: 30057 Tokyo
3: 15573 Japan
4: 870389 Ministry_of_Finance_(Japan)
5: 5405 China
1: 15573 Japan
    [pem_pos=0, final_scores=0.6770754,
final_local_scores=0.65775347, pem=-0.13549899,
similarity=24.455252]
2: 30057 Tokyo
    [pem_pos=0, final_scores=0.66250724,
final_local_scores=0.6825437, pem=-0.09283106,
similarity=24.963007]
3: 15573 Japan
    [pem_pos=0, final_scores=0.70329297,
final_local_scores=0.6863922, pem=-0.13549899,
similarity=25.309889]
4: 870389 Ministry_of_Finance_(Japan)
    [pem_pos=3, final_scores=0.3267042,
final_local_scores=0.41259485, pem=-2.2930157,
similarity=28.87214]
5: 5405 China
    [pem_pos=0, final_scores=0.5154839,
final_local_scores=0.5281922, pem=-0.15669827,
similarity=20.704172]
```

Sample: b'791_PRESS'

PRESS DIGEST - HK newspaper editorials - Aug 28 . [1 1 HONG KONG 1] 1996-08-28 With 307 days to go before the [2 2 British 2] colony reverts to [3 3 China 3] , the [4 4 Hong Kong 4] media focused mainly on domestic issues concerning alleged pressure on a judge , cross straits relations and the democratic lobby 's relationship with [5 5 Beijing 5] . The Beijing-funded [6 6 WEN WEI PO 6] said [7 7 Taiwan 7] 's government could not hope to stem the island 's economic and trade exchanges with [8 8 China 8] . The paper said that using administrative power to limit economic activities across the [9 9 Taiwan 9] strait would not work . [10 10 MING PAO DAILY NEWS 10] said it hoped [11 11 Chinese 11] officials would soon open dialogue with [12 12 Hong Kong 12] 's [13 13 Democratic Party 13] and the newly-established democracy lobby , Frontier , in order to ease anxieties in the lead-up to the handover . The [14 14 English 14] language [15 15 SOUTH CHINA MORNING POST 15] said the judiciary needed to take swift and decisive action in investigating the allegations that a judge had been subjected to pressure in a [16 16 New Zealand 16] immigration case involving allegations of fraud . The independence of the judiciary and the rule of law were of paramount importance to [17 17 Hong Kong 17] 's survival as a business centre . The [18 18 Chinese 18] language daily [19 19 HONG KONG ECONOMIC TIMES 19] said the [20 20 Legal Department 20] had been indecisive in its handling of the judge 's case . Such hesitancy on the part of the government had damaged public confidence in the rule of law , the paper said . -- [21 21 Hong Kong 21] newsroom (852) 2843 6441

1: 13404 Hong_Kong

2: 31717 United_Kingdom

3: 5405 China

4: 13404 Hong_Kong

5: 5405 China

6: 346226 Wen_Wei_Po

7: 25734 Taiwan

8: 5405 China

9: 25734 Taiwan

10: 268276 Ming_Pao

11: 5405 China

12: 13404 Hong_Kong

13: 397335 Democratic_Party_(Hong_Kong)

14: 8569916 English_language

15: 6816735 South_China_Morning_Post

16: 4913064 New_Zealand

17: 13404 Hong_Kong

18: 5751 Chinese_language

19: 348517 Hong_Kong_Economic_Times

20: 2286075 Legal_Department

21: 13404 Hong_Kong

1: 13404 Hong_Kong

```
[pem_pos=0, final_scores=0.55749017,
final_local_scores=0.5070372, pem=-0.22402541,
similarity=20.439001]
2: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.52958816,
final_local_scores=0.5297137, pem=-0.48936236,
similarity=22.558655]
3: 5405 China
    [pem_pos=0, final_scores=0.5202199,
final_local_scores=0.49097443, pem=-0.15669827,
similarity=19.59352]
4: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5051612,
final_local_scores=0.44987583, pem=-0.22402541,
similarity=18.73319]
5: 5405 China
    [pem_pos=7, final_scores=0.22018073,
final_local_scores=0.1632272, pem=-3.4768412,
similarity=27.868341]
6: 346226 Wen_Wei_Po
    [pem_pos=0, final_scores=0.67166656,
final_local_scores=0.7345395, pem=0.0, similarity=26.009836]
7: 25734 Taiwan
    [pem_pos=0, final_scores=0.44693923,
final_local_scores=0.40306115, pem=-0.077886656,
similarity=16.541422]
8: 5405 China
    [pem_pos=0, final_scores=0.53550726,
final_local_scores=0.5076735, pem=-0.15669827,
similarity=20.091852]
9: 25734 Taiwan
    [pem_pos=0, final_scores=0.37570608,
final_local_scores=0.32524976, pem=-0.077886656,
similarity=14.219375]
10: 268276 Ming_Pao
    [pem_pos=0, final_scores=0.43423635,
final_local_scores=0.46594322, pem=0.0, similarity=17.994389]
11: 5405 China
    [pem_pos=1, final_scores=0.73866534,
final_local_scores=0.72959286, pem=-1.4196125,
similarity=33.582294]
12: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.56317884,
final_local_scores=0.51325125, pem=-0.22402541,
similarity=20.62444]
13: 397335 Democratic_Party_(Hong_Kong)
    [pem_pos=5, final_scores=0.11500622,
final_local_scores=0.14278573, pem=-3.6196127,
similarity=28.03474]
14: 8569916 English_language
    [pem_pos=0, final_scores=0.4473831,
final_local_scores=0.47802657, pem=-0.640266,
```

```
similarity=21.836843]
15: 6816735 South_China_Morning_Post
    [pem_pos=0, final_scores=0.37935287,
final_local_scores=0.39176494, pem=0.0, similarity=15.780761]
16: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.5395472,
final_local_scores=0.5663864, pem=-0.23795462,
similarity=22.285847]
17: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5221948,
final_local_scores=0.46848238, pem=-0.22402541,
similarity=19.288446]
18: 5751 Chinese_language
    [pem_pos=0, final_scores=0.12575822,
final_local_scores=0.10067104, pem=-1.1254069,
similarity=13.214075]
19: 348517 Hong_Kong_Economic_Times
    [pem_pos=0, final_scores=0.47669595,
final_local_scores=0.5580892, pem=0.0, similarity=20.744207]
20: 2286075 Legal_Department
    [pem_pos=0, final_scores=0.33622485,
final_local_scores=0.48735112, pem=0.0, similarity=18.633242]
21: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.4795416,
final_local_scores=0.42189026, pem=-0.22402541,
similarity=17.898045]
```

Sample: b'792_Palestinian'

```
[1 1 Palestinian Authority 1] frees rights activist . [2 2 GA
ZA 2] 1996-08-28 A human rights activist said on Wednesday he
had been released after more than two weeks in detention that
followed his call for an inquiry into the death of a [3 3 Gaza
3] man interrogated by [4 4 Palestinian 4] police . Mohammad
Dahman , director of the Gaza-based Addameer Prisoners
Support Association , said he was freed on Tuesday without
being charged . [5 5 Palestinian 5] Attorney-General Khaled
al-Qidra was not immediately available to comment . Qidra had
said Dahman was arrested on suspicion of making a false
statement . The activist was detained by [6 6 Palestinian 6]
intelligence service agents on August 12 after publishing a
statement demanding an investigation into the death of a [7 7
Gaza 7] man who had been questioned by [8 8 Palestinian 8]
authorities . The [9 9 Palestinian Authority 9] said the dead
man , Nahed Dahlan , had committed suicide . Human rights
groups had protested about Dahman 's arrest in letters to [10
10 Palestinian 10] President [11 11 Yasser Arafat 11] and to
Qidra .
```

1: 24093 Palestinian_National_Authority
2: 55387 Gaza
3: 55387 Gaza
4: 241405 State_of_Palestine
5: 241405 State_of_Palestine

```
6: 241405 State_of_Palestine
7: 55387 Gaza
8: 241405 State_of_Palestine
9: 24093 Palestinian_National_Authority
10: 241405 State_of_Palestine
11: 34289 Yasser_Arafat
1: 24093 Palestinian_National_Authority
    [pem_pos=0, final_scores=0.4307781,
final_local_scores=0.37330097, pem=0.0, similarity=15.229758]
2: 55387 Gaza
    [pem_pos=0, final_scores=0.28693563,
final_local_scores=0.24285775, pem=-0.30319342,
similarity=12.985887]
3: 55387 Gaza
    [pem_pos=0, final_scores=0.31123072,
final_local_scores=0.26939648, pem=-0.30319342,
similarity=13.777854]
4: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.6663646,
final_local_scores=0.626824, pem=-2.8508017,
similarity=38.298496]
5: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.5842131,
final_local_scores=0.5370859, pem=-2.8508017,
similarity=35.62053]
6: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.72410434,
final_local_scores=0.68989587, pem=-2.8508017,
similarity=40.180687]
7: 55387 Gaza
    [pem_pos=0, final_scores=0.32171696,
final_local_scores=0.28085107, pem=-0.30319342,
similarity=14.119681]
8: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.7171123,
final_local_scores=0.6822581, pem=-2.8508017,
similarity=39.95276]
9: 24093 Palestinian_National_Authority
    [pem_pos=0, final_scores=0.38110977,
final_local_scores=0.31904587, pem=0.0, similarity=13.610678]
10: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.72411555,
final_local_scores=0.68990815, pem=-2.8508017,
similarity=40.181053]
11: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.58468527,
final_local_scores=0.5788625, pem=0.0, similarity=21.364124]
```

Sample: b'793_PRESS'

PRESS DIGEST - [1 1 Spain 1] - Aug 28 . Headlines from major
national newspapers . [2 2 Reuters 2] has not verified these
stories and does not vouch for their accuracy . [3 3 EL PAIS 3

] - Work groups and weekend arrest to quell juvenile violence
in [4 4 Basque Country 4] [5 5 EL MUNDO 5] - Aleix Vidal-
Quadras - [6 6 Catalan 6] nationalists are demanding my
defenestration [7 7 DIARIO 16 7] - [8 8 Catalan 8]
nationalists say the 1997 budget will make Spaniards sweat [9
9 ABC 9] - Worldwide alarm over child prostitution CINCO DIAS
- [10 10 Banco Santander 10] starts conquest of the east . [11
11 EXPANSION 11] - Government will finish pension reform
before the year 2000 GACETA DE LOS NEGOCIOS - Caja de Madrid
stagnates during struggle for presidency

1: 26667 Spain
2: 18998750 Reuters
3: 444409 El País
4: 20514897 Basque_Country_(autonomous_community)
5: 539568 El_Mundo_(Spain)
6: 6822 Catalonia
7: 6242592 Diario_16
8: 6822 Catalonia
9: 62027 American_Broadcasting_Company
10: 2629368 Santander_Group
11: 6517516 Expansión
1: 26667 Spain
[pem_pos=0, final_scores=0.60777116,
final_local_scores=0.6168687, pem=-0.18857051,
similarity=23.52378]
2: 18998750 Reuters
[pem_pos=0, final_scores=0.30754882,
final_local_scores=0.3637538, pem=0.0, similarity=14.944851]
3: 444409 El País
[pem_pos=0, final_scores=0.41632873,
final_local_scores=0.41589826, pem=0.0, similarity=16.500946]
4: 20514897 Basque_Country_(autonomous_community)
[pem_pos=0, final_scores=0.30365193,
final_local_scores=0.30983496, pem=-0.23203784,
similarity=14.597664]
5: 539568 El_Mundo_(Spain)
[pem_pos=0, final_scores=0.4636268,
final_local_scores=0.4817301, pem=0.0, similarity=18.4655]
6: 6822 Catalonia
[pem_pos=1, final_scores=0.45749646,
final_local_scores=0.44053298, pem=-1.5206007,
similarity=25.50536]
7: 6242592 Diario_16
[pem_pos=0, final_scores=0.45947695,
final_local_scores=0.5809335, pem=0.0, similarity=21.425928]
8: 6822 Catalonia
[pem_pos=1, final_scores=0.44076842,
final_local_scores=0.4222601, pem=-1.5206007,
similarity=24.960058]
9: 62027 American_Broadcasting_Company
[pem_pos=0, final_scores=0.30035067,
final_local_scores=0.39259648, pem=-0.3336378,

```
similarity=17.61995]
10: 2629368 Santander_Group
    [pem_pos=0, final_scores=0.3879369,
final_local_scores=0.4520958, pem=0.0, similarity=17.581154]
11: 6517516 Expansión
    [pem_pos=3, final_scores=0.5071644,
final_local_scores=0.67544657, pem=-1.609438,
similarity=32.99876]

Sample: b'794_Iran'
[1 1 Iran 1] asks [2 2 Bonn 2] to extradite ex-president [3 3
Banisadr 3] . [4 4 BONN 4] 1996-08-28 [5 5 Iran 5] has asked [
6 6 Germany 6] to extradite its former president [7 7 Abolhas
san Banisadr 7] for alleged hijacking , an [8 8 Iranian 8]
embassy spokesman said on Wednesday . [9 9 Banisadr 9]
angered [10 10 Tehran 10] last week by accusing top [11 11 Ir
anian 11] leaders of ordering the assassination of [12 12 Ira
nian Kurdish 12] leaders in a [13 13 Berlin 13] restaurant in
1992 . He made the allegations at the trial of an [14 14 Iran
ian 14] and four [15 15 Lebanese 15] accused of carrying out
the attack . An [16 16 Iranian 16] embassy spokesman said in
response to an inquiry that [17 17 Iran 17] had formally
requested [18 18 Banisadr 18] 's extradition for hijacking
the military aircraft which he commandeered to flee [19 19 Ir
an 19] in July 1981 . " We submitted the request three or
four days ago , " he said . [20 20 German 20] authorities
could not immediately be reached for comment . [21 21 Banisadr
21] lives under round-the-clock security in [22 22 France 22]
, fearing [23 23 Tehran 23] could make an attempt on his
life . He is due back in [24 24 Berlin 24] on September 5 to
continue his testimony , which has backed up [25 25 German 25]
prosecutors ' allegations that [26 26 Tehran 26] ordered the
attack on the exiled leaders . Three dissidents and their
translator were killed in the gangland-style machinegun
attack . [27 27 Iran 27] has warned [28 28 Germany 28] that
bilateral relations could suffer if it pays heed to the
testimony of [29 29 Banisadr 29] , an architect of [30 30 Iran
30] 's [31 31 Islamic 31] revolution who has been a sworn
enemy of [32 32 Tehran 32] since he fell from favour after a
year as president .
1: 14653 Iran
2: 3295 Bonn
3: 399571 Abolhassan_Banisadr
4: 3295 Bonn
5: 14653 Iran
6: 11867 Germany
7: 399571 Abolhassan_Banisadr
8: 14653 Iran
9: 399571 Abolhassan_Banisadr
10: 14653 Iran
11: 14653 Iran
12: 24993749 Iranian_Kurdistan
```

```
13: 3354 Berlin
14: 14653 Iran
15: 17771 Lebanon
16: 14653 Iran
17: 14653 Iran
18: 399571 Abolhassan_Banisadr
19: 14653 Iran
20: 11867 Germany
21: 399571 Abolhassan_Banisadr
22: 5843419 France
23: 14653 Iran
24: 3354 Berlin
25: 11867 Germany
26: 14653 Iran
27: 14653 Iran
28: 11867 Germany
29: 399571 Abolhassan_Banisadr
30: 14653 Iran
31: 6037917 Islam
32: 14653 Iran
1: 14653 Iran
    [pem_pos=0, final_scores=0.63183755,
final_local_scores=0.6149404, pem=-0.13856095,
similarity=23.194277]
2: 3295 Bonn
    [pem_pos=0, final_scores=0.4682908,
final_local_scores=0.5177747, pem=0.0, similarity=19.541143]
3: 399571 Abolhassan_Banisadr
    [pem_pos=0, final_scores=0.48732895,
final_local_scores=0.55272025, pem=0.0, similarity=20.583988]
4: 3295 Bonn
    [pem_pos=0, final_scores=0.38856834,
final_local_scores=0.43069005, pem=0.0, similarity=16.942364]
5: 14653 Iran
    [pem_pos=0, final_scores=0.67274946,
final_local_scores=0.65963054, pem=-0.13856095,
similarity=24.52792]
6: 11867 Germany
    [pem_pos=0, final_scores=0.78448385,
final_local_scores=0.844315, pem=-0.15529114,
similarity=30.130255]
7: 399571 Abolhassan_Banisadr
    [pem_pos=0, final_scores=0.62979287,
final_local_scores=0.70834035, pem=0.0, similarity=25.228003]
8: 14653 Iran
    [pem_pos=0, final_scores=0.6252287,
final_local_scores=0.6077213, pem=-0.57066226,
similarity=25.32868]
9: 399571 Abolhassan_Banisadr
    [pem_pos=0, final_scores=0.53309876,
final_local_scores=0.6027168, pem=0.0, similarity=22.075985]
10: 14653 Iran
```

```
[pem_pos=5, final_scores=0.4283554,
final_local_scores=0.3926671, pem=-2.6998966,
similarity=30.490139]
11: 14653 Iran
    [pem_pos=0, final_scores=0.5896484,
final_local_scores=0.5688553, pem=-0.57066226,
similarity=24.168839]
12: 24993749 Iranian_Kurdistan
    [pem_pos=0, final_scores=0.4209916,
final_local_scores=0.4300995, pem=0.0, similarity=16.92474]
13: 3354 Berlin
    [pem_pos=0, final_scores=0.6544824,
final_local_scores=0.7206284, pem=-0.15483057,
similarity=26.436695]
14: 14653 Iran
    [pem_pos=0, final_scores=0.48454076,
final_local_scores=0.45404112, pem=-0.57066226,
similarity=20.742556]
15: 17771 Lebanon
    [pem_pos=0, final_scores=0.5999914,
final_local_scores=0.6134733, pem=-0.55808944,
similarity=25.431957]
16: 14653 Iran
    [pem_pos=0, final_scores=0.61785203,
final_local_scores=0.59966344, pem=-0.57066226,
similarity=25.088215]
17: 14653 Iran
    [pem_pos=0, final_scores=0.71717614,
final_local_scores=0.7081599, pem=-0.13856095,
similarity=25.976133]
18: 399571 Abolhassan_Banisadr
    [pem_pos=0, final_scores=0.5586434,
final_local_scores=0.63062036, pem=0.0, similarity=22.908682]
19: 14653 Iran
    [pem_pos=0, final_scores=0.5010709,
final_local_scores=0.47209775, pem=-0.13856095,
similarity=18.931568]
20: 11867 Germany
    [pem_pos=1, final_scores=0.6946666,
final_local_scores=0.74620336, pem=-1.0465823,
similarity=32.04939]
21: 399571 Abolhassan_Banisadr
    [pem_pos=0, final_scores=0.5286327,
final_local_scores=0.5978383, pem=0.0, similarity=21.930399]
22: 5843419 France
    [pem_pos=0, final_scores=0.63191503,
final_local_scores=0.67979914, pem=-0.16240834,
similarity=25.259476]
23: 14653 Iran
    [pem_pos=5, final_scores=0.43492007,
final_local_scores=0.39983803, pem=-2.6998966,
similarity=30.704132]
```

```
24: 3354 Berlin
    [pem_pos=0, final_scores=0.47154027,
final_local_scores=0.5207919, pem=-0.15483057,
similarity=20.473175]
25: 11867 Germany
    [pem_pos=1, final_scores=0.7448504,
final_local_scores=0.8010215, pem=-1.0465823,
similarity=33.685272]
26: 14653 Iran
    [pem_pos=5, final_scores=0.46201545,
final_local_scores=0.42943567, pem=-2.6998966,
similarity=31.587385]
27: 14653 Iran
    [pem_pos=0, final_scores=0.6619714,
final_local_scores=0.6478571, pem=-0.13856095,
similarity=24.176579]
28: 11867 Germany
    [pem_pos=0, final_scores=0.7251856,
final_local_scores=0.77954066, pem=-0.15529114,
similarity=28.197258]
29: 399571 Abolhassan_Banisadr
    [pem_pos=0, final_scores=0.5932382,
final_local_scores=0.66841, pem=0.0, similarity=24.0364]
30: 14653 Iran
    [pem_pos=0, final_scores=0.6535886,
final_local_scores=0.6387002, pem=-0.13856095,
similarity=23.903316]
31: 6037917 Islam
    [pem_pos=0, final_scores=0.3966866,
final_local_scores=0.43279368, pem=-0.21141872,
similarity=18.154867]
32: 14653 Iran
    [pem_pos=5, final_scores=0.44942373,
final_local_scores=0.41568106, pem=-2.6998966,
similarity=31.17692]
```

Sample: b'795_NATO'

[1 1 NATO 1] military chiefs to visit [2 2 Iberia 2] . [3 3 B RUSSELS 3] 1996-08-28 Top military officials from [4 4 North Atlantic Treaty Organisation 4] countries will tour [5 5 Spain 5] and [6 6 Portugal 6] next month for their annual inspection of alliance country installations and forces . [7 7 NATO 7] said in a statement received on Wednesday that its military committee would visit the two countries between September 8 and 13 . The committee consists of the chiefs of defence staff of each alliance country except [8 8 Iceland 8] , which has no armed forces . [9 9 NATO 9] 's top military men -- General [10 10 George Joulwan 10] , Supreme Allied Commander [11 11 Europe 11] , and General [12 12 John Sheehan 12] , Supreme Allied Commander Atlantic -- will also attend . The committee 's last tour was in September 1995 in [13 13 Belgium 13] , [14 14 Luxembourg 14] and the [15 15 Netherlands 1

```
5] . REUTER
1: 21133 NATO
2: 14883 Iberian_Peninsula
3: 3708 Brussels
4: 21133 NATO
5: 26667 Spain
6: 23033 Portugal
7: 21133 NATO
8: 14531 Iceland
9: 21133 NATO
10: 774454 George_Joulwan
11: 9239 Europe
12: 8640622 John_J._Sheehan
13: 3343 Belgium
14: 17515 Luxembourg
15: 21148 Netherlands
1: 21133 NATO
    [pem_pos=0, final_scores=0.39954376,
final_local_scores=0.37664002, pem=0.0, similarity=15.329402]
2: 14883 Iberian_Peninsula
    [pem_pos=0, final_scores=0.4056465,
final_local_scores=0.4314677, pem=-0.65314126,
similarity=20.517454]
3: 3708 Brussels
    [pem_pos=0, final_scores=0.55368876,
final_local_scores=0.5500063, pem=-0.025440974,
similarity=20.64135]
4: 21133 NATO
    [pem_pos=0, final_scores=0.43779165,
final_local_scores=0.41842002, pem=0.0, similarity=16.5762]
5: 26667 Spain
    [pem_pos=0, final_scores=0.58280617,
final_local_scores=0.58949506, pem=-0.18857051,
similarity=22.706898]
6: 23033 Portugal
    [pem_pos=0, final_scores=0.5876091,
final_local_scores=0.5952006, pem=-0.30942732,
similarity=23.534399]
7: 21133 NATO
    [pem_pos=0, final_scores=0.40779275,
final_local_scores=0.3856508, pem=0.0, similarity=15.598302]
8: 14531 Iceland
    [pem_pos=0, final_scores=0.56174415,
final_local_scores=0.5772407, pem=-0.36386928,
similarity=23.294504]
9: 21133 NATO
    [pem_pos=0, final_scores=0.48114818,
final_local_scores=0.4657805, pem=0.0, similarity=17.989532]
10: 774454 George_Joulwan
    [pem_pos=0, final_scores=0.38445407,
final_local_scores=0.5339646, pem=0.0, similarity=20.02428]
11: 9239 Europe
```

```
[pem_pos=0, final_scores=0.5748062,
final_local_scores=0.5742793, pem=-0.086338006,
similarity=21.696873]
12: 8640622 John_J._Sheehan
    [pem_pos=0, final_scores=0.33231455,
final_local_scores=0.4505592, pem=-0.050525267,
similarity=17.810062]
13: 3343 Belgium
    [pem_pos=0, final_scores=0.6202368,
final_local_scores=0.60531175, pem=-0.24180195,
similarity=23.46838]
14: 17515 Luxembourg
    [pem_pos=0, final_scores=0.4644146,
final_local_scores=0.4406662, pem=-0.49845144,
similarity=19.950727]
15: 21148 Netherlands
    [pem_pos=0, final_scores=0.38740885,
final_local_scores=0.3591888, pem=-0.24940795,
similarity=16.164942]
```

Sample: b'796_ISS'

ISS says agreed sale of [1 1 U.S. 1] unit . [2 2 COPENHAGEN 2] 1996-08-28 [3 3 Danish 3] cleaning group ISS on Wednesday said it had signed a letter of intent to sell its troubled [4 4 U.S 4] unit ISS Inc to [5 5 Canadian 5] firm Aaxis Limited . An ISS statement said that Aaxis , with year-end 1996 assets of [6 6 US\$ 6] 10.9 million and equity of \$ 10.5 million , would be listed on the [7 7 Montreal 7] stock exchange , but did not say when . It said that under the sale agreement , full financial details of which were not revealed , ISS would acquire a 25 percent stake in Aaxis which would become an associated company within the ISS group trading under the ISS name and logo . ISS Inc senior management would continue to run the business under the new owners , it said . [8 8 Danish 8] analysts recently estimated ISS Inc 's sale value at up to \$ 118 million . ISS said that the deal included ISS Inc operations in [9 9 Mexico 9] and the sale of ISS Inc interests in [10 10 Brazil 10] would be discussed . On August 15 , ISS published first half 1996 results showing a two billion crown loss caused by falsified accounts in ISS Inc and said that charges and provisions earlier estimated at \$ 100 million would have to be increased to \$ 146 million . It also wrote down all ISS Inc goodwill and Wednesday 's statement said that the Aaxis purchase would not necessitate further write down if the sale were completed according to the terms of the letter of intent . -- Steve Weizman , [11 11 Copenhagen 11] newsroom +45 33969650

```
1: 3434750 United_States
2: 5166 Copenhagen
3: 76972 Denmark
4: 3434750 United_States
5: 5042916 Canada
```

```
6: 18717338 United_States_dollar
7: 7954681 Montreal
8: 76972 Denmark
9: 3966054 Mexico
10: 3383 Brazil
11: 5166 Copenhagen
1: 3434750 United_States
    [pem_pos=0, final_scores=0.47056735,
final_local_scores=0.42890853, pem=-0.14075686,
similarity=17.654657]
2: 5166 Copenhagen
    [pem_pos=0, final_scores=0.6885163,
final_local_scores=0.6869901, pem=-0.11866472,
similarity=25.236183]
3: 76972 Denmark
    [pem_pos=0, final_scores=0.6191872,
final_local_scores=0.59328574, pem=-0.65021485,
similarity=25.330513]
4: 3434750 United_States
    [pem_pos=0, final_scores=0.44365525,
final_local_scores=0.39951116, pem=0.0, similarity=16.011923]
5: 5042916 Canada
    [pem_pos=0, final_scores=0.6507676,
final_local_scores=0.6369673, pem=-0.41303828,
similarity=25.344254]
6: 18717338 United_States_dollar
    [pem_pos=0, final_scores=0.48920053,
final_local_scores=0.54697317, pem=0.0, similarity=20.412483]
7: 7954681 Montreal
    [pem_pos=0, final_scores=0.5711437,
final_local_scores=0.5786371, pem=-0.11023275,
similarity=21.956861]
8: 76972 Denmark
    [pem_pos=0, final_scores=0.51505524,
final_local_scores=0.47953743, pem=-0.65021485,
similarity=21.936037]
9: 3966054 Mexico
    [pem_pos=0, final_scores=0.6998524,
final_local_scores=0.7218456, pem=-0.18258883,
similarity=26.623972]
10: 3383 Brazil
    [pem_pos=0, final_scores=0.7834571,
final_local_scores=0.8198712, pem=-0.21313342,
similarity=29.715357]
11: 5166 Copenhagen
    [pem_pos=0, final_scores=0.56220907,
final_local_scores=0.5490185, pem=-0.11866472,
similarity=21.118837]
```

Sample: b'797_Iraq'
[1 1 Iraq 1] balks at [2 2 U.N. 2] staff for oil-for-food
deal . Evelyn Leopold [3 3 UNITED NATIONS 3] 1996-08-28 [4 4 I

Iraq 4] has balked at the number of [5 5 U.N. 5] staff needed to implement the oil-for-food deal , blaming the [6 6 United States 6] for insisting on stringent monitoring . In comments to reporters and a statement on Tuesday , [7 7 Iraqi 7] diplomats said the cost of the monitors and other staff , which Baghdad has to finance , surpasses funds allocated for electricity , water , sewers , education and agriculture . At issue was a May 20 agreement allowing [8 8 Iraq 8] to sell \$ 2 billion worth of oil to purchase badly needed food , medicine and other supplies to ease the impact of sanctions in force since its troops invaded [9 9 Kuwait 9] in August 1990 . The [10 10 Iraqi 10] statement said the [11 11 United States 11] was " interfering and pressing to augment the number of international staff and this is not legal and not justified . " [12 12 Iraq 12] 's deputy ambassador , Saeed Hasan , noted that the May 20 accord said that the number of personnel would be determined by the [13 13 United Nations 13] and that the government of [14 14 Iraq 14] would be consulted . Saeed in his comments did not threaten to call off the deal and the [15 15 U.N. 15] officials said they expected it to go into force next month after Secretary-General [16 16 Boutros Boutros-Ghali 16] reports that arrangements are in place . The U.N. Department of Humanitarian Affairs (DHA) , which has to coordinate the distribution of food , medicine and other goods , increased the number of monitors earlier this month at the insistence of the [17 17 United States 17] . According to [18 18 U.N. 18] officials and diplomats , [19 19 Iraq 19] would have about \$ 1.13 billion to spend for food , medicine and other goods after monies for a reparations fund for [20 20 Gulf War 20] victims and costs for [21 21 U.N. 21] weapons inspections were deducted . The cost of the [22 22 U.N. 22] staff overseeing the distribution of food and other supplies was estimated to cost \$ 31 million . In addition another \$ 12 million was anticipated to cover other expenses , such as oil experts and administrative costs . For the distribution and supervision of humanitarian supplies the [23 23 United Nations 23] estimated it needed 1,190 people , including 267 international staff and 923 [24 24 Iraqi 24] support staff . Of this number 64 foreign and 598 local staff would be in the northern [25 25 Kurdish 25] provinces , no longer the direct control of the [26 26 Baghdad 26] government . Another 203 international staff and 325 [27 27 Iraqis 27] would run the programme in the central and southern parts of the country . There are also 14 monitors to watch oil flows , 32 customs experts and four [28 28 New York-based 28] oil experts or overseers to approve contracts . [29 29 Yasushi Akashi 29] , the DHA undersecretary-general , told the [30 30 Security Council 30] last week that the " financial requirements to support the (humanitarian) programme represent a very modest percentile of the total ... roughly 3 percent . " 1: 7515928 Iraq

```
2: 31769 United_Nations
3: 31769 United_Nations
4: 7515928 Iraq
5: 31769 United_Nations
6: 3434750 United_States
7: 7515928 Iraq
8: 7515928 Iraq
9: 7515890 Kuwait
10: 7515928 Iraq
11: 3434750 United_States
12: 7515928 Iraq
13: 31769 United_Nations
14: 7515928 Iraq
15: 31769 United_Nations
16: 211193 Boutros_Boutros-Ghali
17: 3434750 United_States
18: 31769 United_Nations
19: 7515928 Iraq
20: 182000 Gulf_War
21: 31769 United_Nations
22: 31769 United_Nations
23: 31769 United_Nations
24: 7515928 Iraq
25: 80777 Kurdistan
26: 7515928 Iraq
27: 7515928 Iraq
28: 645042 New_York_City
29: 1271160 Yasushi_Akashi
30: 31956 United_Nations_Security_Council
1: 7515928 Iraq
    [pem_pos=0, final_scores=0.59651333,
final_local_scores=0.55471224, pem=-0.25379938,
similarity=22.023634]
2: 31769 United_Nations
    [pem_pos=0, final_scores=0.53305644,
final_local_scores=0.495076, pem=0.0, similarity=18.863768]
3: 31769 United_Nations
    [pem_pos=0, final_scores=0.43773562,
final_local_scores=0.39095247, pem=0.0, similarity=15.756514]
4: 7515928 Iraq
    [pem_pos=0, final_scores=0.62013173,
final_local_scores=0.58051175, pem=-0.25379938,
similarity=22.793543]
5: 31769 United_Nations
    [pem_pos=0, final_scores=0.55573595,
final_local_scores=0.51984996, pem=0.0, similarity=19.603073]
6: 3434750 United_States
    [pem_pos=0, final_scores=0.46694225,
final_local_scores=0.46666634, pem=-0.073457584,
similarity=18.415442]
7: 7515928 Iraq
    [pem_pos=0, final_scores=0.5209951,
```

```
final_local_scores=0.47222006, pem=-0.48469657,
similarity=20.817556]
8: 7515928 Iraq
    [pem_pos=0, final_scores=0.5697383,
final_local_scores=0.5254646, pem=-0.25379938,
similarity=21.150826]
9: 7515890 Kuwait
    [pem_pos=0, final_scores=0.43507206,
final_local_scores=0.39736444, pem=-0.3372397,
similarity=17.781822]
10: 7515928 Iraq
    [pem_pos=0, final_scores=0.50240505,
final_local_scores=0.45191324, pem=-0.48469657,
similarity=20.211561]
11: 3434750 United_States
    [pem_pos=0, final_scores=0.5018543,
final_local_scores=0.5048024, pem=-0.073457584,
similarity=19.553497]
12: 7515928 Iraq
    [pem_pos=0, final_scores=0.57981235,
final_local_scores=0.5364689, pem=-0.25379938,
similarity=21.479216]
13: 31769 United_Nations
    [pem_pos=0, final_scores=0.45987463,
final_local_scores=0.41513604, pem=0.0, similarity=16.4782]
14: 7515928 Iraq
    [pem_pos=0, final_scores=0.6377085,
final_local_scores=0.5997117, pem=-0.25379938,
similarity=23.366507]
15: 31769 United_Nations
    [pem_pos=0, final_scores=0.56862724,
final_local_scores=0.53393173, pem=0.0, similarity=20.0233]
16: 211193 Boutros_Boutros-Ghali
    [pem_pos=0, final_scores=0.6682449,
final_local_scores=0.7149581, pem=0.0, similarity=25.425488]
17: 3434750 United_States
    [pem_pos=0, final_scores=0.48356843,
final_local_scores=0.48482788, pem=-0.073457584,
similarity=18.957418]
18: 31769 United_Nations
    [pem_pos=0, final_scores=0.5264711,
final_local_scores=0.4878825, pem=0.0, similarity=18.6491]
19: 7515928 Iraq
    [pem_pos=0, final_scores=0.60863626,
final_local_scores=0.5679547, pem=-0.25379938,
similarity=22.418816]
20: 182000 Gulf_War
    [pem_pos=0, final_scores=0.352606,
final_local_scores=0.3268333, pem=0.0, similarity=13.843071]
21: 31769 United_Nations
    [pem_pos=0, final_scores=0.5349539,
final_local_scores=0.49714875, pem=0.0, similarity=18.925623]
```

```
22: 31769 United_Nations
    [pem_pos=0, final_scores=0.54428035,
final_local_scores=0.5073365, pem=0.0, similarity=19.229647]
23: 31769 United_Nations
    [pem_pos=0, final_scores=0.49397558,
final_local_scores=0.45238614, pem=0.0, similarity=17.589817]
24: 7515928 Iraq
    [pem_pos=0, final_scores=0.54801124,
final_local_scores=0.5017311, pem=-0.48469657,
similarity=21.698225]
25: 80777 Kurdistan
    [pem_pos=9, final_scores=0.34803945,
final_local_scores=0.33630574, pem=-2.8001814,
similarity=29.353569]
26: 7515928 Iraq
    [pem_pos=8, final_scores=0.21803688,
final_local_scores=0.14128423, pem=-2.5644467,
similarity=22.251776]
27: 7515928 Iraq
    [pem_pos=1, final_scores=0.53050065,
final_local_scores=0.48260343, pem=-0.8129937,
similarity=22.912748]
28: 645042 New_York_City
    [pem_pos=1, final_scores=0.40976173,
final_local_scores=0.46078283, pem=-0.062242296,
similarity=18.178875]
29: 1271160 Yasushi_Akashi
    [pem_pos=0, final_scores=0.7912827,
final_local_scores=0.91249555, pem=0.0, similarity=31.3204]
30: 31956 United_Nations_Security_Council
    [pem_pos=0, final_scores=0.41246122,
final_local_scores=0.40213037, pem=0.0, similarity=16.090086]
```

Sample: b'798_OFFICIAL'

OFFICIAL JOURNAL CONTENTS - OJ C 251 OF AUGUST 29 , 1996 . *

(Note - contents are displayed in reverse order to that in the printed Journal) * Aircraft noise and emissions Economic assessment of proposals for a common [1 1 European Union 1] position for CAEP 4 Consultancy services Call for tender (96 / C 251/09) Provision of overland transport services for material and equipment for [2 2 European Commission 2] delegations in European Third Countries and in the New Independent States (NIS) Contract notice No TRA / 96 / 003 / IAE-3 - Open procedure (96 / C 251/08) Microfiche production system Open procedure Invitation to tender DI 96/04 Micromation (96 / C 251/07) Aircraft noise and emissions Gaseous emissions from aircraft in the atmosphere Consultancy services Call for tender (96 / C 251/06) Tacis - support framework for the coordination and development of the Tacis information and communications programme Notice of open invitation to tender for a public service contract (96 / C 251/05) COUNCIL REGULATION ([3 3 EEC 3]) No

4064/89 (96 / C 251/04) FINANCIAL STATEMENTS OF THE [4 4 EU
ROPEAN 4] COAL AND STEEL COMMUNITY AT 31 DECEMBER 1995 (96 /
C 251/03) Average prices and representative prices for table
wines at the various marketing centres (96 / C 251/02) Ecu
(1) 28 August 1996 (96 / C 251/01) END OF DOCUMENT .

1: 9317 European_Union
2: 9974 European_Commission
3: 9578 European_Economic_Community
4: 9239 Europe
1: 9317 European_Union
[pem_pos=0, final_scores=0.5246187,
final_local_scores=0.4726817, pem=-0.005814796,
similarity=18.2271]
2: 9974 European_Commission
[pem_pos=0, final_scores=0.52685744,
final_local_scores=0.50988215, pem=0.0, similarity=19.305613]
3: 9578 European_Economic_Community
[pem_pos=0, final_scores=0.2785251,
final_local_scores=0.23722556, pem=0.0, similarity=11.168997]
4: 9239 Europe
[pem_pos=0, final_scores=0.50498676,
final_local_scores=0.5508135, pem=-0.25557294,
similarity=21.916931]

Sample: b'799_EU'
[1 1 EU Commission 1] cool on changing beef cull plan . [2 2
BRUSSELS 2] 1996-08-29 The [3 3 European Commission 3] said
on Thursday it would study scientific reports saying [4 4 Bri
tain 4] 's mad cow epidemic would die out by 2001 but offered
little prospect the findings would change an agreed slaughter
campaign . " Obviously we are interested in this research .
We will ask the (EU) scientific and veterinary committee to
examine it , " [5 5 Commission 5] spokesman Gerard Kiely told
[6 6 Reuters 6] . But he added that new research into the
dynamics of the bovine spongiform encephalopathy ([7 7 BSE 7]
), a fatal brain-wasting disease suffered by cattle , was
unlikely to alter a slaughter plan agreed by [8 8 Britain 8]
and its 14 EU partners . " We agreed that following detailed
scientific analysis using a methodology which would take out
the maximum number of [9 9 BSE 9] cases possible . I think it
would be very difficult to sell to the [10 10 European Commis
sion 10] a programme which would involve the elimination of
fewer [11 11 BSE 11] cases , " Kiely said . " We will look at
our approach (to the plan) but we wo n't get involved with
the number of animals to be slaughtered , " he said . " We
have always avoided the question of numbers of animals to be
slaughtered , that 's not the issue . The issue is the
protection of consumers ' health and the rapid eradication of
[12 12 BSE 12] , " he added . The reaction is likely to
disappoint [13 13 British 13] farmers , who seized on
research by [14 14 Oxford 14] scientists in the scientific
journal [15 15 Nature 15] saying it would be hard to get rid

of the disease any faster than 2001 without killing vast numbers of cattle . The researchers predicted there would be 340 new infections and 14,000 new cases of [16 16 BSE 16] before 2001 . [17 17 British 17] farmers ' leader called on Wednesday for an urgent meeting with ministers to discuss the report . " I hope the government will now make it clear they believe there is a better way of dealing with this issue , " National Farmers Union president Sir David Naish told BBC radio . Naish said there was no need for [18 18 Britain 18] to carry out a planned cull of some 147,000 cattle to which it had reluctantly agreed to placate its [19 19 European 19] partners . " The new evidence to me means some of that proposal should be re-examined because we could get away with considerably less animals being culled if in fact scientists throughout [20 20 Europe 20] accepted this evidence , " Naish said . The report could well reopen a damaging row between [21 21 Britain 21] and the EU , which slapped a worldwide ban on [22 22 British 22] beef after the government said there could be a link between [23 23 BSE 23] and the human form of the disease . The issue flared in in March when government scientists admitted that people could become infected with [24 24 Creutzfeldt-Jakob Disease 24] ([25 25 CJD 25]) from eating BSE-infected beef .

1: 9974 European_Commission
2: 3708 Brussels
3: 9974 European_Commission
4: 31717 United_Kingdom
5: 9974 European_Commission
6: 18998750 Reuters
7: 19344418 Bovine_spongiform_encephalopathy
8: 31717 United_Kingdom
9: 19344418 Bovine_spongiform_encephalopathy
10: 9974 European_Commission
11: 19344418 Bovine_spongiform_encephalopathy
12: 19344418 Bovine_spongiform_encephalopathy
13: 31717 United_Kingdom
14: 31797 University_of_Oxford
15: 43427 Nature_(journal)
16: 19344418 Bovine_spongiform_encephalopathy
17: 31717 United_Kingdom
18: 31717 United_Kingdom
19: 9317 European_Union
20: 9239 Europe
21: 31717 United_Kingdom
22: 31717 United_Kingdom
23: 19344418 Bovine_spongiform_encephalopathy
24: 5617 Creutzfeldt-Jakob_disease
25: 5617 Creutzfeldt-Jakob_disease
1: 9974 European_Commission
[pem_pos=0, final_scores=0.44055408,
final_local_scores=0.43169987, pem=0.0, similarity=16.972498]
2: 3708 Brussels

```
[pem_pos=0, final_scores=0.48989218,
final_local_scores=0.5182954, pem=-0.025440974,
similarity=19.695034]
3: 9974 European_Commission
    [pem_pos=0, final_scores=0.48478514,
final_local_scores=0.48001552, pem=0.0, similarity=18.414333]
4: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5314955,
final_local_scores=0.5058785, pem=-0.61114657,
similarity=22.509645]
5: 9974 European_Commission
    [pem_pos=0, final_scores=0.531317,
final_local_scores=0.5308445, pem=-0.76247066,
similarity=24.077606]
6: 18998750 Reuters
    [pem_pos=0, final_scores=0.38412797,
final_local_scores=0.42106634, pem=0.0, similarity=16.655172]
7: 19344418 Bovine_spongiform_encephalopathy
    [pem_pos=0, final_scores=0.6950386,
final_local_scores=0.7261196, pem=-0.2201173,
similarity=26.9556]
8: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5647387,
final_local_scores=0.54219157, pem=-0.61114657,
similarity=23.5933]
9: 19344418 Bovine_spongiform_encephalopathy
    [pem_pos=0, final_scores=0.72830695,
final_local_scores=0.7624601, pem=-0.2201173,
similarity=28.040075]
10: 9974 European_Commission
    [pem_pos=0, final_scores=0.44267797,
final_local_scores=0.43401986, pem=0.0, similarity=17.04173]
11: 19344418 Bovine_spongiform_encephalopathy
    [pem_pos=0, final_scores=0.7281986,
final_local_scores=0.7623418, pem=-0.2201173,
similarity=28.036545]
12: 19344418 Bovine_spongiform_encephalopathy
    [pem_pos=0, final_scores=0.7229915,
final_local_scores=0.7566538, pem=-0.2201173,
similarity=27.866802]
13: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5739903,
final_local_scores=0.55229753, pem=-0.48936236,
similarity=23.232601]
14: 31797 University_of_Oxford
    [pem_pos=1, final_scores=0.2637598,
final_local_scores=0.29782897, pem=-1.4798834,
similarity=21.02536]
15: 43427 Nature_(journal)
    [pem_pos=0, final_scores=0.4248007,
final_local_scores=0.47808194, pem=-0.4031761,
similarity=20.549166]
```

```
16: 19344418 Bovine_spongiform_encephalopathy
    [pem_pos=0, final_scores=0.7033879,
final_local_scores=0.7352398, pem=-0.2201173,
similarity=27.227766]
17: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.49594915,
final_local_scores=0.46704948, pem=-0.48936236,
similarity=20.68863]
18: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5584694,
final_local_scores=0.53534335, pem=-0.61114657,
similarity=23.388935]
19: 9317 European_Union
    [pem_pos=1, final_scores=0.21618365,
final_local_scores=0.18159717, pem=-2.9373715,
similarity=25.482819]
20: 9239 Europe
    [pem_pos=0, final_scores=0.57958037,
final_local_scores=0.60146165, pem=-0.086338006,
similarity=22.508049]
21: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5505741,
final_local_scores=0.5267189, pem=-0.61114657,
similarity=23.131565]
22: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.4883923,
final_local_scores=0.4587947, pem=-0.48936236,
similarity=20.442291]
23: 19344418 Bovine_spongiform_encephalopathy
    [pem_pos=0, final_scores=0.73101336,
final_local_scores=0.76541644, pem=-0.2201173,
similarity=28.128298]
24: 5617 Creutzfeldt-Jakob_disease
    [pem_pos=0, final_scores=0.49794,
final_local_scores=0.5668219, pem=0.0, similarity=21.004808]
25: 5617 Creutzfeldt-Jakob_disease
    [pem_pos=0, final_scores=0.5123475,
final_local_scores=0.5825599, pem=0.0, similarity=21.474463]
```

Sample: b'800_French'

```
[1 1 French 1] farmers set up blockades in mad cow protest .
[2 2 PARIS 2] 1996-08-29 Thousands of farmers threw up
roadblocks across [3 3 France 3] overnight , stopping and
checking lorries suspected of importing meat from outside the
[4 4 European Union 4] , [5 5 French 5] radios reported on
Thursday . Radio stations said around 15,000 farmers ,
angered by a fall in beef prices following the mad cow
disease crisis , staged protests in many areas and blockaded
several main roads and motorways . By 3 a.m. ( 0100 [6 6 GMT 6
] ) more than 2,000 lorries had been stopped and searched . [7
7 European 7] beef sales plunged after [8 8 Britain 8]
announced the discovery of a likely link between bovine
```

spongiform encephalopathy ([9 9 BSE 9]) , or mad cow disease , and its fatal human equivalent [10 10 Creutzfeldt-Jakob Disease 10] ([11 11 CJD 11]) .

1: 5843419 France
2: 22989 Paris
3: 5843419 France
4: 9317 European_Union
5: 5843419 France
6: 12701 Greenwich_Mean_Time
7: 9239 Europe
8: 31717 United_Kingdom
9: 19344418 Bovine_spongiform_encephalopathy
10: 5617 Creutzfeldt-Jakob_disease
11: 5617 Creutzfeldt-Jakob_disease

1: 5843419 France
[pem_pos=0, final_scores=0.72143424,
final_local_scores=0.71052814, pem=-0.80223274,
similarity=29.655956]
2: 22989 Paris
[pem_pos=0, final_scores=0.4903288,
final_local_scores=0.51039594, pem=-0.115301244,
similarity=19.947971]
3: 5843419 France
[pem_pos=0, final_scores=0.660193,
final_local_scores=0.64363146, pem=-0.16240834,
similarity=24.18016]
4: 9317 European_Union
[pem_pos=0, final_scores=0.41881698,
final_local_scores=0.43128842, pem=-0.005814796,
similarity=16.991842]
5: 5843419 France
[pem_pos=0, final_scores=0.6009478,
final_local_scores=0.578915, pem=-0.80223274,
similarity=25.728357]
6: 12701 Greenwich_Mean_Time
[pem_pos=0, final_scores=0.6587632,
final_local_scores=0.7750978, pem=0.0, similarity=27.220177]
7: 9239 Europe
[pem_pos=0, final_scores=0.39633197,
final_local_scores=0.38001466, pem=-0.25557294,
similarity=16.819954]
8: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5512294,
final_local_scores=0.54117054, pem=-0.61114657,
similarity=23.562832]
9: 19344418 Bovine_spongiform_encephalopathy
[pem_pos=0, final_scores=0.7283776,
final_local_scores=0.8225171, pem=-0.2201173,
similarity=29.832294]
10: 5617 Creutzfeldt-Jakob_disease
[pem_pos=0, final_scores=0.5298135,
final_local_scores=0.62561524, pem=0.0, similarity=22.75932]

```
11: 5617 Creutzfeldt-Jakob_disease
    [pem_pos=0, final_scores=0.55656433,
final_local_scores=0.6548365, pem=0.0, similarity=23.63134]
```

Sample: b'801_GOLF'

GOLF - [1 1 BRITISH MASTERS 1] SECOND ROUND SCORES . [2 2 NO RTHAMPTON 2] , [3 3 England 3] 1996-08-29 Leading scores after the second round of the [4 4 British Masters 4] on Thursday ([5 5 British 5] unless stated) : 140 [6 6 Robert A llenby 6] ([7 7 Australia 7]) 69 71 , [8 8 Mark Roe 8] 69 71 141 Francisco Cea ([9 9 Spain 9]) 70 71 , Gavin Levenson ([10 10 South Africa 10]) 66 75 142 [11 11 Daniel Chopra 11] ([12 12 Sweden 12]) 74 68 143 [13 13 David Gilford 13] 69 74 144 Peter O'Malley ([14 14 Australia 14]) 71 73 , [15 15 Costantino Rocca 15] ([16 16 Italy 16]) 71 73 , [17 17 Colin Montgomerie 17] 68 76 , [18 18 David Howell 18] 70 74 , [19 19 Mark Davis 19] 71 73 145 [20 20 Peter Mitchell 20] 74 71 , [21 21 Philip Walton 21] ([22 22 Ireland 22]) 71 74 , [23 23 Re tief Goosen 23] ([24 24 South Africa 24]) 71 74 , [25 25 Ove Sellberg 25] ([26 26 Sweden 26]) 71 74 , [27 27 Peter Hedbl om 27] ([28 28 Sweden 28]) 70 75 , [29 29 Pedro Linhart 29] ([30 30 Spain 30]) 72 73 , Mike Clayton ([31 31 Australia 3 1]) 69 76 , [32 32 Emanuele Canonica 32] ([33 33 Italy 33]) 69 76 , [34 34 Miguel Angel Martin 34] ([35 35 Spain 35]) 75 70 146 [36 36 Iain Pyman 36] 71 75 , [37 37 Eduardo Ro mero 37] ([38 38 Argentina 38]) 70 76 , [39 39 Ian Woosnam 3 9] 70 76 , [40 40 Miguel Angel Jimenez 40] ([41 41 Spain 41]) 74 72 , [42 42 Klas Eriksson 42] ([43 43 Sweden 43]) 71 75 , [44 44 Paul Eales 44] 75 71 147 Antoine Lebouc ([45 45 France 45]) 74 73 , [46 46 Paul Curry 46] 76 71 , [47 47 Andr ew Coltart 47] 72 75 , [48 48 Paul Lawrie 48] 72 75 , [49 49 Jose Caceres 49] ([50 50 Argentina 50]) 69 78 , [51 51 Raymo nd Russell 51] 69 78 , [52 52 Roger Chapman 52] 71 76 , [53 53 Paul Affleck 53] 74 73 .

1: 1929870 British_Masters
2: 63734 Northampton
3: 9316 England
4: 1929870 British_Masters
5: 31717 United_Kingdom
6: 1983254 Robert_Allenby
7: 4689264 Australia
8: 4353394 Mark_Roe
9: 26667 Spain
10: 17416221 South_Africa
11: 2371253 Daniel_Chopra
12: 5058739 Sweden
13: 2028738 David_Gilford
14: 4689264 Australia
15: 2157151 Costantino_Rocca
16: 14532 Italy
17: 379616 Colin_Montgomerie
18: 1343807 David_Howell_(golfer)

```
19: 5120820 Mark_Davis_(golfer)
20: 5000237 Peter_Mitchell_(golfer)
21: 5179238 Philip_Walton
22: 14560 Republic_of_Ireland
23: 630525 Retief_Goosen
24: 17416221 South_Africa
25: 5177483 Ove_Sellberg
26: 5058739 Sweden
27: 4377532 Peter_Hedblom
28: 5058739 Sweden
29: 5285076 Pedro_Linhart
30: 26667 Spain
31: 4689264 Australia
32: 2401223 Emanuele_Canonica
33: 14532 Italy
34: 5022198 Miguel_Ángel_Martín_(golfer)
35: 26667 Spain
36: 5088382 Iain_Pyman
37: 2119368 Eduardo_Romero
38: 18951905 Argentina
39: 1034991 Ian_Woosnam
40: 1569625 Miguel_Ángel_Jiménez
41: 26667 Spain
42: 28186295 Klas_Eriksson
43: 5058739 Sweden
44: 4435993 Paul_Eales
45: 5843419 France
46: 5340359 Paul_Curry_(golfer)
47: 2150917 Andrew_Coltart
48: 1564436 Paul_Lawrie
49: 2156041 José_Cóceres
50: 18951905 Argentina
51: 4375906 Raymond_Russell
52: 4380250 Roger_Chapman_(golfer)
53: 24997138 Paul_Affleck
1: 1929870 British_Masters
    [pem_pos=0, final_scores=0.40539807,
final_local_scores=0.46880108, pem=0.0, similarity=18.079672]
2: 63734 Northampton
    [pem_pos=0, final_scores=0.4173931,
final_local_scores=0.47733587, pem=-0.5110059,
similarity=21.113297]
3: 9316 England
    [pem_pos=0, final_scores=0.55673975,
final_local_scores=0.5898718, pem=-0.19053856,
similarity=22.728844]
4: 1929870 British_Masters
    [pem_pos=0, final_scores=0.41344088,
final_local_scores=0.47758663, pem=0.0, similarity=18.34185]
5: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.50311583,
final_local_scores=0.5449106, pem=-0.48936236,
```

```
similarity=23.012161]
6: 1983254 Robert_Allenby
    [pem_pos=0, final_scores=0.7373306,
final_local_scores=0.8115742, pem=0.0, similarity=28.308706]
7: 4689264 Australia
    [pem_pos=0, final_scores=0.646764,
final_local_scores=0.69998235, pem=-0.17168477,
similarity=25.912231]
8: 4353394 Mark_Roe
    [pem_pos=0, final_scores=0.72437125,
final_local_scores=0.8097605, pem=0.0, similarity=28.254581]
9: 26667 Spain
    [pem_pos=0, final_scores=0.6035408,
final_local_scores=0.62975055, pem=-0.18857051,
similarity=23.908201]
10: 17416221 South_Africa
    [pem_pos=0, final_scores=0.5465389,
final_local_scores=0.6035217, pem=-0.26251382,
similarity=23.527596]
11: 2371253 Daniel_Chopra
    [pem_pos=0, final_scores=0.5207976,
final_local_scores=0.6093095, pem=0.0, similarity=22.272724]
12: 5058739 Sweden
    [pem_pos=0, final_scores=0.6772526,
final_local_scores=0.7137024, pem=-0.2135508,
similarity=26.549337]
13: 2028738 David_Gilford
    [pem_pos=0, final_scores=0.73286563,
final_local_scores=0.8248268, pem=0.0, similarity=28.70419]
14: 4689264 Australia
    [pem_pos=0, final_scores=0.65290886,
final_local_scores=0.70669466, pem=-0.17168477,
similarity=26.112541]
15: 2157151 Costantino_Rocca
    [pem_pos=0, final_scores=0.6724244,
final_local_scores=0.7319077, pem=0.0, similarity=25.9313]
16: 14532 Italy
    [pem_pos=0, final_scores=0.7533675,
final_local_scores=0.82077324, pem=-0.20612453,
similarity=29.70416]
17: 379616 Colin_Montgomerie
    [pem_pos=0, final_scores=0.705476,
final_local_scores=0.74259335, pem=0.0, similarity=26.25018]
18: 1343807 David_Howell_(golfer)
    [pem_pos=0, final_scores=0.554858,
final_local_scores=0.60046583, pem=-0.78602827,
similarity=26.283354]
19: 5120820 Mark_Davis_(golfer)
    [pem_pos=8, final_scores=0.4588424,
final_local_scores=0.52074444, pem=-2.4415264,
similarity=32.907166]
20: 5000237 Peter_Mitchell_(golfer)
```

```
[pem_pos=3, final_scores=0.5169621,
final_local_scores=0.57572657, pem=-1.2033675,
similarity=27.814642]
21: 5179238 Philip_Walton
    [pem_pos=0, final_scores=0.83721703,
final_local_scores=0.90183944, pem=0.0, similarity=31.002401]
22: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.57176584,
final_local_scores=0.6253733, pem=-1.5040284,
similarity=30.93124]
23: 630525 Retief_Goosen
    [pem_pos=0, final_scores=0.7536132,
final_local_scores=0.832142, pem=0.0, similarity=28.92249]
24: 17416221 South_Africa
    [pem_pos=0, final_scores=0.57286984,
final_local_scores=0.6322842, pem=-0.26251382,
similarity=24.385927]
25: 5177483 Ove_Sellberg
    [pem_pos=0, final_scores=0.5966656,
final_local_scores=0.671696, pem=0.0, similarity=24.134462]
26: 5058739 Sweden
    [pem_pos=0, final_scores=0.6831208,
final_local_scores=0.72011256, pem=-0.2135508,
similarity=26.74063]
27: 4377532 Peter_Hedblom
    [pem_pos=0, final_scores=0.6119023,
final_local_scores=0.67425036, pem=0.0, similarity=24.21069]
28: 5058739 Sweden
    [pem_pos=0, final_scores=0.751646,
final_local_scores=0.7949659, pem=-0.2135508,
similarity=28.974403]
29: 5285076 Pedro_Linhart
    [pem_pos=0, final_scores=0.7140803,
final_local_scores=0.8030017, pem=0.0, similarity=28.052885]
30: 26667 Spain
    [pem_pos=0, final_scores=0.6116715,
final_local_scores=0.6386321, pem=-0.18857051,
similarity=24.173244]
31: 4689264 Australia
    [pem_pos=0, final_scores=0.63943523,
final_local_scores=0.69197685, pem=-0.17168477,
similarity=25.673332]
32: 2401223 Emanuele_Canonica
    [pem_pos=0, final_scores=0.6384603,
final_local_scores=0.7461097, pem=0.0, similarity=26.355116]
33: 14532 Italy
    [pem_pos=0, final_scores=0.7414597,
final_local_scores=0.8077657, pem=-0.20612453,
similarity=29.31599]
34: 5022198 Miguel_Ángel_Martín_(golfer)
    [pem_pos=0, final_scores=0.66161877,
final_local_scores=0.72750825, pem=0.0, similarity=25.80001]
```

35: 26667 Spain
[pem_pos=0, final_scores=0.5826633,
final_local_scores=0.60694504, pem=-0.18857051,
similarity=23.227638]
36: 5088382 Iain_Pyman
[pem_pos=0, final_scores=0.6337941,
final_local_scores=0.7109851, pem=0.0, similarity=25.306927]
37: 2119368 Eduardo_Romero
[pem_pos=0, final_scores=0.6315698,
final_local_scores=0.68169445, pem=0.0, similarity=24.432835]
38: 18951905 Argentina
[pem_pos=0, final_scores=0.66468245,
final_local_scores=0.7096009, pem=-0.30003807,
similarity=26.897274]
39: 1034991 Ian_Woosnam
[pem_pos=0, final_scores=0.69833046,
final_local_scores=0.7393904, pem=0.0, similarity=26.154596]
40: 1569625 Miguel_Ángel_Jiménez
[pem_pos=0, final_scores=0.44571722,
final_local_scores=0.488621, pem=0.0, similarity=18.671139]
41: 26667 Spain
[pem_pos=0, final_scores=0.58463925,
final_local_scores=0.6091035, pem=-0.18857051,
similarity=23.292051]
42: 28186295 Klas_Eriksson
[pem_pos=0, final_scores=0.60295504,
final_local_scores=0.6978631, pem=0.0, similarity=24.91534]
43: 5058739 Sweden
[pem_pos=0, final_scores=0.67482954,
final_local_scores=0.71105564, pem=-0.2135508,
similarity=26.470354]
44: 4435993 Paul_Eales
[pem_pos=0, final_scores=0.6513698,
final_local_scores=0.72302216, pem=0.0, similarity=25.666136]
45: 5843419 France
[pem_pos=0, final_scores=0.7382777,
final_local_scores=0.7927917, pem=-0.16240834,
similarity=28.631401]
46: 5340359 Paul_Curry_(golfer)
[pem_pos=1, final_scores=0.5905016,
final_local_scores=0.6744211, pem=-0.5474354,
similarity=27.192822]
47: 2150917 Andrew_Coltart
[pem_pos=0, final_scores=0.64873964,
final_local_scores=0.6959573, pem=0.0, similarity=24.858467]
48: 1564436 Paul_Lawrie
[pem_pos=0, final_scores=0.8542078,
final_local_scores=0.9300042, pem=0.0, similarity=31.842892]
49: 2156041 José_Cóceres
[pem_pos=0, final_scores=0.650394,
final_local_scores=0.70523536, pem=0.0, similarity=25.135344]
50: 18951905 Argentina

```
[pem_pos=0, final_scores=0.6452784,
final_local_scores=0.688405, pem=-0.30003807,
similarity=26.264744]
51: 4375906 Raymond_Russell
    [pem_pos=0, final_scores=0.61662585,
final_local_scores=0.67376024, pem=0.0, similarity=24.196062]
52: 4380250 Roger_Chapman_(golfer)
    [pem_pos=1, final_scores=0.64745075,
final_local_scores=0.705916, pem=-0.5573004,
similarity=28.18634]
53: 24997138 Paul_Affleck
    [pem_pos=0, final_scores=0.34842026,
final_local_scores=0.45215034, pem=0.0, similarity=17.58278]
```

Sample: b'802_CYCLING'

CYCLING - SORENSEN WINS FOURTH STAGE OF TOUR OF
NETHERLANDS . [1 1 DOETINCHEM 1] , [2 2 Netherlands 2]
1996-08-29 Leading results and overall standings after the
19.6 kilometre fourth stage of the [3 3 Tour of the Netherlan
ds 3] on Thursday , a time trial starting and finishing in [4
4 Doetinchem 4] . 1. [5 5 Rolf Sorensen 5] ([6 6 Denmark 6])
[7 7 Rabobank 7] 22 minutes 40 seconds 2. [8 8 Lance Armstro
ng 8] ([9 9 U.S. 9]) [10 10 Motorola 10] 1 second behind 3.
[11 11 Vyacheslav Ekimov 11] ([12 12 Russia 12]) [13 13 Rab
obank 13] 29 seconds behind 4. [14 14 Erik Dekker 14] ([15 15
Netherlands 15]) [16 16 Rabobank 16] 43 5. Giunluca Gorini ([17 17 Italy 17]) Aki 45 6. [18 18 Erik Breukink 18] ([19 19
Netherlands 19]) [20 20 Rabobank 20] 48 7. [21 21 Wilfried Pe
eters 21] ([22 22 Belgium 22]) Mapei 51 8. [23 23 Bart Vosk
amp 23] ([24 24 Netherlands 24]) [25 25 TVM 25] 53 9. [26 26
Michael Andersson 26] ([27 27 Sweden 27]) Telekom 54 10.
Gregory Randolph ([28 28 USA 28]) [29 29 Motorola 29] 1
minute 3 seconds Leading overall placings after three
stages : 1. Sorensen 11.20:33 2. [30 30 Armstrong 30] 3
seconds behind 3. [31 31 Ekimov 31] 1 minute 7 seconds 4. [32
32 Marco Lietti 32] ([33 33 Italy 33]) MG-Technogym 1
minute 14 seconds 5. [34 34 Dekker 34] 1 minute 21 seconds 6.
[35 35 Breukink 35] 1 minute 26 seconds 7. [36 36 Maarten den
Bakker 36] ([37 37 Netherlands 37]) [38 38 TVM 38] 1 minute
31 seconds 8. [39 39 Voskamp 39] same time 9. [40 40 Andersson
40] 1 minute 32 seconds 10. [41 41 Olaf Ludwig 41] ([42 42 Ge
rmany 42]) Telekom 1 minute 44 seconds The race continues on
Friday with the 178 kilometre fifth-stage from [43 43 Zevenaar
43] to [44 44 Venray 44] .

1: 118600 Doetinchem
2: 21148 Netherlands
3: 11065220 Ronde_van_Nederland
4: 118600 Doetinchem
5: 2611279 Rolf_Sørensen
6: 76972 Denmark
7: 2354465 Belkin_Pro_Cycling_Team
8: 23243880 Lance_Armstrong

```
9: 3434750 United_States
10: 2290911 7-Eleven_(cycling_team)
11: 914182 Viatcheslav_Ekimov
12: 25391 Russia
13: 2354465 Belkin_Pro_Cycling_Team
14: 934261 Erik_Dekker
15: 21148 Netherlands
16: 2354465 Belkin_Pro_Cycling_Team
17: 14532 Italy
18: 2234080 Erik_Breukink
19: 21148 Netherlands
20: 2354465 Belkin_Pro_Cycling_Team
21: 10711801 Wilfried_Peeters
22: 3343 Belgium
23: 11066143 Bart_Voskamp
24: 21148 Netherlands
25: 2642474 TVM_(cycling_team)
26: 11100004 Michael_Andersson_(cyclist)
27: 5058739 Sweden
28: 3434750 United_States
29: 2290911 7-Eleven_(cycling_team)
30: 23243880 Lance_Armstrong
31: 914182 Viatcheslav_Ekimov
32: 11099892 Marco_Lietti
33: 14532 Italy
34: 934261 Erik_Dekker
35: 2234080 Erik_Breukink
36: 11100389 Maarten_den_Bakker
37: 21148 Netherlands
38: 2642474 TVM_(cycling_team)
39: 11066143 Bart_Voskamp
40: 11100004 Michael_Andersson_(cyclist)
41: 2107369 Olaf_Ludwig
42: 11867 Germany
43: 118651 Zevenaar
44: 118556 Venray
1: 118600 Doetinchem
    [pem_pos=0, final_scores=0.42257714,
final_local_scores=0.48043644, pem=0.0, similarity=18.426895]
2: 21148 Netherlands
    [pem_pos=2, final_scores=0.53590363,
final_local_scores=0.5390443, pem=-0.24940795,
similarity=21.53219]
3: 11065220 Ronde_van_Nederland
    [pem_pos=0, final_scores=0.5964407,
final_local_scores=0.6704016, pem=0.0, similarity=24.095833]
4: 118600 Doetinchem
    [pem_pos=0, final_scores=0.3956867,
final_local_scores=0.45106268, pem=0.0, similarity=17.550323]
5: 2611279 Rolf_Sørensen
    [pem_pos=0, final_scores=0.7026238,
final_local_scores=0.77485293, pem=0.0, similarity=27.21287]
```

6: 76972 Denmark
[pem_pos=0, final_scores=0.59203047,
final_local_scores=0.634111, pem=-0.26762357,
similarity=24.468227]
7: 2354465 Belkin_Pro_Cycling_Team
[pem_pos=0, final_scores=0.5251531,
final_local_scores=0.5409807, pem=0.0, similarity=20.233656]
8: 23243880 Lance_Armstrong
[pem_pos=0, final_scores=0.60306984,
final_local_scores=0.6405796, pem=0.0, similarity=23.205885]
9: 3434750 United_States
[pem_pos=0, final_scores=0.4233703,
final_local_scores=0.48021537, pem=-0.14075686,
similarity=19.185757]
10: 2290911 7-Eleven_(cycling_team)
[pem_pos=1, final_scores=0.66606367,
final_local_scores=0.7404051, pem=-1.7647687,
similarity=35.781963]
11: 914182 Viatcheslav_Ekimov
[pem_pos=0, final_scores=0.7753946,
final_local_scores=0.85170037, pem=0.0, similarity=29.506151]
12: 25391 Russia
[pem_pos=0, final_scores=0.71132725,
final_local_scores=0.81915915, pem=-0.20299491,
similarity=29.638973]
13: 2354465 Belkin_Pro_Cycling_Team
[pem_pos=0, final_scores=0.58116263,
final_local_scores=0.60216254, pem=0.0, similarity=22.059444]
14: 934261 Erik_Dekker
[pem_pos=0, final_scores=0.8356387,
final_local_scores=0.90406996, pem=0.0, similarity=31.068962]
15: 21148 Netherlands
[pem_pos=2, final_scores=0.64669424,
final_local_scores=0.66006625, pem=-0.24940795,
similarity=25.143726]
16: 2354465 Belkin_Pro_Cycling_Team
[pem_pos=0, final_scores=0.5136808,
final_local_scores=0.52844894, pem=0.0, similarity=19.859684]
17: 14532 Italy
[pem_pos=0, final_scores=0.72149795,
final_local_scores=0.7904646, pem=-0.20612453,
similarity=28.79969]
18: 2234080 Erik_Breukink
[pem_pos=0, final_scores=0.9893351,
final_local_scores=1.0616233, pem=0.0, similarity=35.77067]
19: 21148 Netherlands
[pem_pos=2, final_scores=0.6092534,
final_local_scores=0.6191678, pem=-0.24940795,
similarity=23.923235]
20: 2354465 Belkin_Pro_Cycling_Team
[pem_pos=0, final_scores=0.53031915,
final_local_scores=0.5466238, pem=0.0, similarity=20.402058]

21: 10711801 Wilfried_Peeters
[pem_pos=0, final_scores=0.68889487,
final_local_scores=0.7727592, pem=0.0, similarity=27.150389]
22: 3343 Belgium
[pem_pos=0, final_scores=0.72856563,
final_local_scores=0.7452988, pem=-0.24180195,
similarity=27.645872]
23: 11066143 Bart_Voskamp
[pem_pos=0, final_scores=0.75083035,
final_local_scores=0.8581673, pem=0.0, similarity=29.699137]
24: 21148 Netherlands
[pem_pos=2, final_scores=0.5617846,
final_local_scores=0.5673154, pem=-0.24940795,
similarity=22.375856]
25: 2642474 TVM_(cycling_team)
[pem_pos=1, final_scores=0.7012153,
final_local_scores=0.8220053, pem=-0.84397006,
similarity=33.20963]
26: 11100004 Michael_Andersson_(cyclist)
[pem_pos=1, final_scores=0.316994,
final_local_scores=0.43020666, pem=-0.926762,
similarity=21.967813]
27: 5058739 Sweden
[pem_pos=0, final_scores=0.71641773,
final_local_scores=0.7718428, pem=-0.2135508,
similarity=28.284363]
28: 3434750 United_States
[pem_pos=0, final_scores=0.47813237,
final_local_scores=0.5400346, pem=-0.13038418,
similarity=20.914473]
29: 2290911 7-Eleven_(cycling_team)
[pem_pos=1, final_scores=0.6178205,
final_local_scores=0.6877067, pem=-1.7647687,
similarity=34.20934]
30: 23243880 Lance_Armstrong
[pem_pos=0, final_scores=1.2054393,
final_local_scores=1.3045528, pem=0.0, similarity=43.020172]
31: 914182 Viatcheslav_Ekimov
[pem_pos=0, final_scores=0.67223257,
final_local_scores=0.74596167, pem=0.0, similarity=26.350697]
32: 11099892 Marco_Lietti
[pem_pos=0, final_scores=0.84617144,
final_local_scores=0.9547176, pem=0.0, similarity=32.58039]
33: 14532 Italy
[pem_pos=0, final_scores=0.7120039,
final_local_scores=0.7800938, pem=-0.20612453,
similarity=28.490206]
34: 934261 Erik_Dekker
[pem_pos=0, final_scores=0.89803344,
final_local_scores=0.9786158, pem=0.0, similarity=33.29356]
35: 2234080 Erik_Breukink
[pem_pos=0, final_scores=0.69822675,

```
final_local_scores=0.7436317, pem=0.0, similarity=26.281166]
36: 11100389 Maarten_den_Bakker
    [pem_pos=0, final_scores=0.5472219,
final_local_scores=0.6566526, pem=0.0, similarity=23.685537]
37: 21148 Netherlands
    [pem_pos=2, final_scores=0.5483383,
final_local_scores=0.5526273, pem=-0.24940795,
similarity=21.937534]
38: 2642474 TVM_(cycling_team)
    [pem_pos=1, final_scores=0.6497411,
final_local_scores=0.76577747, pem=-0.84397006,
similarity=31.531681]
39: 11066143 Bart_Voskamp
    [pem_pos=0, final_scores=0.6350596,
final_local_scores=0.73991203, pem=0.0, similarity=26.170164]
40: 11100004 Michael_Andersson_(cyclist)
    [pem_pos=1, final_scores=0.4117394,
final_local_scores=0.53370154, pem=-0.926762,
similarity=25.056307]
41: 2107369 Olaf_Ludwig
    [pem_pos=0, final_scores=0.6260581,
final_local_scores=0.6981852, pem=0.0, similarity=24.924952]
42: 11867 Germany
    [pem_pos=0, final_scores=0.85513216,
final_local_scores=0.91308457, pem=-0.15529114,
similarity=32.182476]
43: 118651 Zevenaar
    [pem_pos=0, final_scores=0.45199245,
final_local_scores=0.5670712, pem=0.0, similarity=21.012249]
44: 118556 Venray
    [pem_pos=0, final_scores=0.43605775,
final_local_scores=0.49904686, pem=0.0, similarity=18.982267]
```

Sample: b'803_CRICKET'

```
CRICKET - [1 1 ENGLAND 1] BEAT [2 2 PAKISTAN 2] IN FIRST ONE-DAYER . [3 3 MANCHESTER 3] , [4 4 England 4] 1996-08-29 [5 5 England 5] beat [6 6 Pakistan 6] by five wickets to win the first one-day ( 50 overs-a-side ) international at [7 7 Old Trafford 7] on Thursday . Scores : [8 8 Pakistan 8] 225-5 innnings closed ( [9 9 Saeed Anwar 9] 57 ) , [10 10 England 10] 226-5 in 46.4 overs ( M. Atherton 65 ) .
```

```
1: 759125 England_cricket_team
2: 759295 Pakistan_national_cricket_team
3: 20206 Manchester
4: 9316 England
5: 759125 England_cricket_team
6: 759295 Pakistan_national_cricket_team
7: 660597 Old_Trafford_Cricket_Ground
8: 759295 Pakistan_national_cricket_team
9: 894437 Saeed_Anwar
10: 759125 England_cricket_team
1: 759125 England_cricket_team
```

```
[pem_pos=2, final_scores=0.70044416,
final_local_scores=0.6559301, pem=-3.3054674,
similarity=41.63962]
2: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.81197625,
final_local_scores=0.7917369, pem=-2.7556365,
similarity=42.7023]
3: 20206 Manchester
    [pem_pos=0, final_scores=0.5199172,
final_local_scores=0.5732169, pem=-0.3209542,
similarity=22.941048]
4: 9316 England
    [pem_pos=0, final_scores=0.57100767,
final_local_scores=0.6228986, pem=-0.19053856,
similarity=23.714428]
5: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.7022806,
final_local_scores=0.65793604, pem=-3.3054674,
similarity=41.699482]
6: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.77899534,
final_local_scores=0.7557102, pem=-2.7556365,
similarity=41.62719]
7: 660597 Old_Trafford_Cricket_Ground
    [pem_pos=1, final_scores=0.5955862,
final_local_scores=0.58812976, pem=-0.6437659,
similarity=25.141577]
8: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.79818887,
final_local_scores=0.77667624, pem=-2.7556365,
similarity=42.25286]
9: 894437 Saeed_Anwar
    [pem_pos=0, final_scores=0.69017154,
final_local_scores=0.7197979, pem=0.0, similarity=25.569918]
10: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.81972426,
final_local_scores=0.7862254, pem=-3.3054674,
similarity=45.527893]
```

Sample: b'804_CYCLING'

```
CYCLING - WORLD CHAMPIONSHIP RESULTS . [1 1 MANCHESTER 1] , [
2 2 England 2] 1996-08-29 Results at the world track cycling
championships on Thursday : Individual pursuit semifinals
( over 4,000 metres ) : [3 3 Chris Boardman 3] ( [4 4 Britain
4] ) 4:15.006 beat [5 5 Alexei Markov 5] ( [6 6 Russia 6] )
4:23.029 [7 7 Andrea Collinelli 7] ( [8 8 Italy 8] ) 4:16.141
beat [9 9 Francis Moreau 9] ( [10 10 France 10] ) 4:19.665 [11
11 Moreau 11] takes bronze medal as faster losing
semifinalist . Final : [12 12 Chris Boardman 12] ( [13 13 Bri
tain 13] ) 4:11.114 ( world record ) beat [14 14 Andrea Colli
nelli 14] ( [15 15 Italy 15] ) 4:20.341 1 Olympic sprint
championship ( three-man teams ) : 1. [16 16 Australia 16] ( [
```

17 17 Darryn Hill 17] , [18 18 Shane Kelly 18] , [19 19 Gary N
eiwand 19]) 44.804 seconds 2. [20 20 Germany 20] ([21 21 Je
ns Fiedler 21] , [22 22 Michael Hubner 22] , [23 23 Soren Lau
sberg 23]) 45.455 3. [24 24 France 24] ([25 25 Laurent Gane
25] , [26 26 Florian Rousseau 26] , Herve Thuet) 45.810 4. [2
7 27 Greece 27] (Dimitrios Georgalis , Georgios Chimonetos ,
Lampros Vasilopoulos) 46.538 Women 's world sprint
championship quarter-finals (best of three matches) :
Magali Faure ([28 28 France 28]) beat Kathrin Freitag ([29
29 Germany 29]) two matches to nil (with times for the last
200 metres of 11.833 seconds and 12.033 seconds) [30 30 Felicia Ballanger 30] ([31 31 France 31]) beat 2 Oksana Grichina ([32 32 Russia 32]) 2-0 , (11.776 / 12.442) [33 33 Tanya Dubnicoff 33] ([34 34 Canada 34]) beat Michelle Ferris ([35
35 Australia 35]) 2-0 , (12.211 / 12.208) [36 36 Annett Neumann 36] ([37 37 Germany 37]) beat Galina Enioukhina ([38
38 Russia 38]) 2-0 , (12.434 / 12.177).

1: 20206 Manchester
2: 9316 England
3: 917770 Chris_Boardman
4: 31717 United_Kingdom
5: 12752327 Alexei_Markov
6: 25391 Russia
7: 17826490 Andrea_Collinelli
8: 14532 Italy
9: 18894901 Francis_Moreau
10: 5843419 France
11: 18894901 Francis_Moreau
12: 917770 Chris_Boardman
13: 31717 United_Kingdom
14: 17826490 Andrea_Collinelli
15: 14532 Italy
16: 4689264 Australia
17: 23203050 Darryn_Hill
18: 4167427 Shane_Kelly
19: 11438632 Gary_Neiwand
20: 11867 Germany
21: 6473320 Jens_Fiedler_(cyclist)
22: 19109147 Michael_Hübner
23: 6691361 Sören_Lausberg
24: 5843419 France
25: 16732280 Laurent_Gané
26: 11438472 Florian_Rousseau
27: 12108 Greece
28: 5843419 France
29: 11867 Germany
30: 6519486 Félicia_Ballanger
31: 5843419 France
32: 25391 Russia
33: 11451462 Tanya_Dubnicoff
34: 5042916 Canada
35: 4689264 Australia

36: 6691583 Annett_Neumann
37: 11867 Germany
38: 25391 Russia
1: 650360 Olympic_sports
[winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.1481288, final_local_scores=-0.19937843,
pem=-3.2532303, similarity=15.831435]
RECALL MISS: 650360 Olympic_sports]
2: 6691515 Oksana_Grishina_(cyclist)
[winner: 27994657 Oksana_Grigorieva, pem_pos=0,
final_scores=-1.216581, final_local_scores=-1.2019813,
pem=-0.6652973, similarity=-28.161818]
RECALL MISS: 6691515 Oksana_Grishina_(cyclist)]
1: 20206 Manchester
[pem_pos=0, final_scores=0.5571651,
final_local_scores=0.57676476, pem=-0.3209542,
similarity=23.046923]
2: 9316 England
[pem_pos=0, final_scores=0.6431231,
final_local_scores=0.6639763, pem=-0.19053856,
similarity=24.94027]
3: 917770 Chris_Boardman
[pem_pos=0, final_scores=0.80137426,
final_local_scores=0.9002679, pem=0.0, similarity=30.955502]
4: 31717 United_Kingdom
[pem_pos=0, final_scores=0.58088094,
final_local_scores=0.5866515, pem=-0.61114657,
similarity=24.920073]
5: 12752327 Alexei_Markov
[pem_pos=0, final_scores=0.5582391,
final_local_scores=0.7080021, pem=0.0, similarity=25.217907]
6: 25391 Russia
[pem_pos=0, final_scores=0.79708225,
final_local_scores=0.8423062, pem=-0.20299491,
similarity=30.329727]
7: 17826490 Andrea_Collinelli
[pem_pos=0, final_scores=0.62566733,
final_local_scores=0.7401012, pem=0.0, similarity=26.17581]
8: 14532 Italy
[pem_pos=0, final_scores=0.77670115,
final_local_scores=0.8016757, pem=-0.20612453,
similarity=29.134253]
9: 18894901 Francis_Moreau
[pem_pos=0, final_scores=0.64481455,
final_local_scores=0.75558907, pem=0.0, similarity=26.637999]
10: 5843419 France
[pem_pos=0, final_scores=0.75367934,
final_local_scores=0.756463, pem=-0.16240834,
similarity=27.54728]
11: 18894901 Francis_Moreau
[pem_pos=0, final_scores=0.5986753,
final_local_scores=0.705189, pem=0.0, similarity=25.133959]

12: 917770 Chris_Boardman
[pem_pos=0, final_scores=0.7941856,
final_local_scores=0.89241534, pem=0.0, similarity=30.721169]
13: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5694085,
final_local_scores=0.5741195, pem=-0.61114657,
similarity=24.546093]
14: 17826490 Andrea_Collinelli
[pem_pos=0, final_scores=0.5923137,
final_local_scores=0.7036674, pem=0.0, similarity=25.088552]
15: 14532 Italy
[pem_pos=0, final_scores=0.7799781,
final_local_scores=0.8052552, pem=-0.20612453,
similarity=29.241072]
16: 4689264 Australia
[pem_pos=0, final_scores=0.57813364,
final_local_scores=0.5934576, pem=-0.17168477,
similarity=22.73332]
17: 23203050 Darryn_Hill
[pem_pos=0, final_scores=0.3897457,
final_local_scores=0.52049315, pem=0.0, similarity=19.622267]
18: 4167427 Shane_Kelly
[pem_pos=0, final_scores=0.41606998,
final_local_scores=0.51653314, pem=0.0, similarity=19.504091]
19: 11438632 Gary_Neiwand
[pem_pos=0, final_scores=0.60356635,
final_local_scores=0.7255213, pem=0.0, similarity=25.740717]
20: 11867 Germany
[pem_pos=0, final_scores=0.7539981,
final_local_scores=0.76288986, pem=-0.15529114,
similarity=27.700365]
21: 6473320 Jens_Fiedler_(cyclist)
[pem_pos=1, final_scores=0.43162304,
final_local_scores=0.5601574, pem=-0.1218049,
similarity=21.46832]
22: 19109147 Michael_Hübner
[pem_pos=0, final_scores=0.5608601,
final_local_scores=0.69192505, pem=0.0, similarity=24.738136]
23: 6691361 Sören_Lausberg
[pem_pos=0, final_scores=0.63097274,
final_local_scores=0.7915171, pem=0.0, similarity=27.710161]
24: 5843419 France
[pem_pos=0, final_scores=0.4807561,
final_local_scores=0.45833576, pem=-0.16240834,
similarity=18.650568]
25: 16732280 Laurent_Gané
[pem_pos=0, final_scores=0.52041525,
final_local_scores=0.67128533, pem=0.0, similarity=24.122206]
26: 11438472 Florian_Rousseau
[pem_pos=0, final_scores=0.49328464,
final_local_scores=0.6119648, pem=0.0, similarity=22.351963]
27: 12108 Greece

```
[pem_pos=0, final_scores=0.721034,
final_local_scores=0.75468767, pem=-0.27607638,
similarity=28.112444]
28: 5843419 France
[pem_pos=0, final_scores=0.773676,
final_local_scores=0.77830625, pem=-0.16240834,
similarity=28.199125]
29: 11867 Germany
[pem_pos=0, final_scores=0.8462903,
final_local_scores=0.863705, pem=-0.15529114,
similarity=30.708889]
30: 6519486 Félicia_Ballanger
[pem_pos=0, final_scores=0.79580337,
final_local_scores=0.94337887, pem=0.0, similarity=32.24202]
31: 5843419 France
[pem_pos=0, final_scores=0.7898304,
final_local_scores=0.7959525, pem=-0.16240834,
similarity=28.725723]
32: 25391 Russia
[pem_pos=0, final_scores=0.81565446,
final_local_scores=0.8625936, pem=-0.20299491,
similarity=30.935143]
33: 11451462 Tanya_Dubnicoff
[pem_pos=0, final_scores=0.7758507,
final_local_scores=0.91068524, pem=0.0, similarity=31.266378]
34: 5042916 Canada
[pem_pos=0, final_scores=0.82485646,
final_local_scores=0.8502852, pem=-0.11662793,
similarity=30.098158]
35: 4689264 Australia
[pem_pos=0, final_scores=0.82762057,
final_local_scores=0.86598414, pem=-0.17168477,
similarity=30.866053]
36: 6691583 Annett_Neumann
[pem_pos=0, final_scores=0.6169971,
final_local_scores=0.7567034, pem=0.0, similarity=26.671251]
37: 11867 Germany
[pem_pos=0, final_scores=0.9141732,
final_local_scores=0.93785685, pem=-0.15529114,
similarity=32.92173]
38: 25391 Russia
[pem_pos=0, final_scores=0.88339025,
final_local_scores=0.93658465, pem=-0.20299491,
similarity=33.143185]
```

Sample: b'805_CRICKET'

CRICKET - [1 1 CROFT 1] RESTRICTS [2 2 PAKISTAN 2] TO
225-5 . [3 3 MANCHESTER 3] , [4 4 England 4] 1996-08-29 Tight
bowling from [5 5 Glamorgan 5] off-spinner [6 6 Robert Croft 6
] helped [7 7 England 7] to restrict [8 8 Pakistan 8] to 225
for five in their 50 overs in the first one-day international
at [9 9 Old Trafford 9] on Thursday . [10 10 Croft 10] , who

was one of the few [11 11 Englishmen 11] to make a good impression in his test debut at [12 12 The Oval 12] last week , showed great control as he first dried up the early flow of [13 13 Pakistan 13] runs and then collected the wickets of Aamir Sohail and [14 14 Wasim Akram 14] in a spell of 10-1-36-2 . There was also a wicket each for [15 15 Ronnie Irani 15] , Allan Mullally and [16 16 Darren Gough 16] although there was no joy for [17 17 Dean Headley 17] who , along with [18 18 Lancashire 18] batsman [19 19 Graham Lloyd 19] , was making his international debut . After [20 20 Wasim 20] had won the toss and chosen to bat first , [21 21 Pakistani 21] made an excellent start as [22 22 Sohail 22] and [23 23 Saeed Anwar 23] continued their good form with an opening partnership of 82 . [24 24 Anwar 24] , who struck a superb 176 at [25 25 The Oval 25] , was the more aggressive as he made 57 from 75 balls before skying a catch off [26 26 Irani 26] to [27 27 Mullally 27] at long-on . [28 28 Sohail 28] and [29 29 Ijaz Ahmed 29] then added 59 for the second wicket before [30 30 England 30] struck back with three wickets for 19 in the space of five overs . First , [31 31 Sohail 31] , after making 48 , was bowled by [32 32 Croft 32] as he stepped back to try and hit through the off-side . [33 33 Wasim 33] , who promoted himself to number four in the order , followed for six when [34 34 Croft 34] drifted another well-flighted delivery behind his legs . Shortly after Ijaz was also back in the pavilion for 48 after [35 35 Irani 35] had repaid [36 36 Mullally 36] with another good catch at long-on . [37 37 Gough 37] later bowled [38 38 Moin Khan 38] with an inswinging yorker but [39 39 Inzamam-ul-Haq 39] , 37 not out , and [40 40 Salim Malik 40] took [41 41 Pakistan 41] to 225 for five when the overs ran out .

- 1: 1160680 Robert_Croft
- 2: 759295 Pakistan_national_cricket_team
- 3: 20206 Manchester
- 4: 9316 England
- 5: 1622320 Glamorgan_County_Cricket_Club
- 6: 1160680 Robert_Croft
- 7: 759125 England_cricket_team
- 8: 759295 Pakistan_national_cricket_team
- 9: 660597 Old_Trafford_Cricket_Ground
- 10: 1160680 Robert_Croft
- 11: 759125 England_cricket_team
- 12: 387687 The_Oval
- 13: 759295 Pakistan_national_cricket_team
- 14: 501393 Wasim_Akram
- 15: 2053953 Ronnie_Irani
- 16: 413178 Darren_Gough
- 17: 672220 Dean_Headley
- 18: 749919 Lancashire_County_Cricket_Club
- 19: 3750666 Graham_Lloyd
- 20: 501393 Wasim_Akram
- 21: 759295 Pakistan_national_cricket_team

```
22: 1436501 Aamer_Sohail
23: 894437 Saeed_Anwar
24: 894437 Saeed_Anwar
25: 387687 The_Oval
26: 2053953 Ronnie_Irani
27: 2222543 Alan_Mullally
28: 1436501 Aamer_Sohail
29: 938863 Ijaz_Ahmed_(cricketer)
30: 759125 England_cricket_team
31: 1436501 Aamer_Sohail
32: 1160680 Robert_Croft
33: 501393 Wasim_Akram
34: 1160680 Robert_Croft
35: 2053953 Ronnie_Irani
36: 2222543 Alan_Mullally
37: 413178 Darren_Gough
38: 554132 Moin_Khan
39: 896403 Inzamam-ul-Haq
40: 938861 Saleem_Malik
41: 759295 Pakistan_national_cricket_team
1: 1160680 Robert_Croft
    [pem_pos=0, final_scores=1.1905336,
final_local_scores=1.2402444, pem=0.0, similarity=41.10108]
2: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.79862684,
final_local_scores=0.7870417, pem=-2.7556365,
similarity=42.562187]
3: 20206 Manchester
    [pem_pos=0, final_scores=0.46436483,
final_local_scores=0.5557353, pem=-0.3209542,
similarity=22.419361]
4: 9316 England
    [pem_pos=0, final_scores=0.52745736,
final_local_scores=0.62533224, pem=-0.19053856,
similarity=23.787052]
5: 1622320 Glamorgan_County_Cricket_Club
    [pem_pos=1, final_scores=0.78728145,
final_local_scores=0.8116551, pem=-0.7903395,
similarity=32.609108]
6: 1160680 Robert_Croft
    [pem_pos=0, final_scores=1.1005019,
final_local_scores=1.1390693, pem=0.0, similarity=38.08181]
7: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.85199517,
final_local_scores=0.83790094, pem=-3.3054674,
similarity=47.069992]
8: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.8349646,
final_local_scores=0.8267352, pem=-2.7556365,
similarity=43.74672]
9: 660597 Old_Trafford_Cricket_Ground
    [pem_pos=1, final_scores=0.60465014,
```

```
final_local_scores=0.594966, pem=-0.6437659,
similarity=25.345585]
10: 1160680 Robert_Croft
    [pem_pos=0, final_scores=1.3549042,
final_local_scores=1.4197941, pem=0.0, similarity=46.4592]
11: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.72127205,
final_local_scores=0.69510573, pem=-1.088303,
similarity=30.751417]
12: 387687 The_Oval
    [pem_pos=0, final_scores=0.6291476,
final_local_scores=0.62259394, pem=0.0, similarity=22.669157]
13: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7094435,
final_local_scores=0.6896226, pem=-2.7556365,
similarity=39.655006]
14: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.76209754,
final_local_scores=0.7572524, pem=0.0, similarity=26.687635]
15: 2053953 Ronnie_Irani
    [pem_pos=0, final_scores=0.72421,
final_local_scores=0.739223, pem=0.0, similarity=26.149601]
16: 413178 Darren_Gough
    [pem_pos=0, final_scores=0.92238265,
final_local_scores=0.9609813, pem=0.0, similarity=32.76731]
17: 672220 Dean_Headley
    [pem_pos=0, final_scores=0.51440215,
final_local_scores=0.585447, pem=0.0, similarity=21.56062]
18: 749919 Lancashire_County_Cricket_Club
    [pem_pos=1, final_scores=0.60791135,
final_local_scores=0.6188613, pem=-1.7295884,
similarity=31.963541]
19: 3750666 Graham_Lloyd
    [pem_pos=0, final_scores=0.48702496,
final_local_scores=0.5500806, pem=0.0, similarity=20.505215]
20: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.9280949,
final_local_scores=0.9409217, pem=0.0, similarity=32.168694]
21: 759295 Pakistan_national_cricket_team
    [pem_pos=2, final_scores=0.75924057,
final_local_scores=0.7440183, pem=-2.3804295,
similarity=39.23785]
22: 1436501 Aamer_Sohail
    [pem_pos=0, final_scores=1.0105368,
final_local_scores=1.0665095, pem=0.0, similarity=35.916485]
23: 894437 Saeed_Anwar
    [pem_pos=0, final_scores=0.9426982,
final_local_scores=0.9671276, pem=0.0, similarity=32.95073]
24: 894437 Saeed_Anwar
    [pem_pos=0, final_scores=1.2199086,
final_local_scores=1.2727842, pem=0.0, similarity=42.072132]
25: 387687 The_Oval
```

```
[pem_pos=0, final_scores=0.61312336,
final_local_scores=0.60508996, pem=0.0, similarity=22.146805]
26: 2053953 Ronnie_Irani
    [pem_pos=0, final_scores=1.16933,
final_local_scores=1.2287095, pem=0.0, similarity=40.75685]
27: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.82106537,
final_local_scores=0.8541639, pem=-0.73930055,
similarity=33.600098]
28: 1436501 Aamer_Sohail
    [pem_pos=0, final_scores=0.8693488,
final_local_scores=0.912283, pem=0.0, similarity=31.314056]
29: 938863 Ijaz_Ahmed_(cricketer)
    [pem_pos=0, final_scores=0.6683092,
final_local_scores=0.6790697, pem=0.0, similarity=24.354507]
30: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.8263324,
final_local_scores=0.80986816, pem=-3.3054674,
similarity=46.23344]
31: 1436501 Aamer_Sohail
    [pem_pos=0, final_scores=0.98776895,
final_local_scores=1.0416391, pem=0.0, similarity=35.1743]
32: 1160680 Robert_Croft
    [pem_pos=0, final_scores=1.4715757,
final_local_scores=1.54724, pem=0.0, similarity=50.262444]
33: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.89136916,
final_local_scores=0.90080434, pem=0.0, similarity=30.971512]
34: 1160680 Robert_Croft
    [pem_pos=0, final_scores=1.3938909,
final_local_scores=1.4623811, pem=0.0, similarity=47.730083]
35: 2053953 Ronnie_Irani
    [pem_pos=0, final_scores=1.0673327,
final_local_scores=1.1172929, pem=0.0, similarity=37.43196]
36: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.7828811,
final_local_scores=0.81245327, pem=-0.73930055,
similarity=32.35537]
37: 413178 Darren_Gough
    [pem_pos=0, final_scores=1.2683938,
final_local_scores=1.3424044, pem=0.0, similarity=44.149734]
38: 554132 Moin_Khan
    [pem_pos=0, final_scores=0.7501968,
final_local_scores=0.8039881, pem=0.0, similarity=28.082321]
39: 896403 Inzamam-ul-Haq
    [pem_pos=0, final_scores=1.0086482,
final_local_scores=1.0368752, pem=0.0, similarity=35.03214]
40: 938861 Saleem_Malik
    [pem_pos=0, final_scores=0.8204014,
final_local_scores=0.862073, pem=0.0, similarity=29.815691]
41: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.84113234,
```

```
final_local_scores=0.83347255, pem=-2.7556365,  
similarity=43.947777]
```

Sample: b'806_CRICKET'
CRICKET - [1 1 ENGLAND 1] V [2 2 PAKISTAN 2] ONE-DAY
SCOREBOARD . [3 3 MANCHESTER 3] , [4 4 England 4] 1996-08-29
Scoreboard of the first one-day (50 overs-a-side) match
between [5 5 England 5] and [6 6 Pakistan 6] at [7 7 Old Traf-
ford 7] on Thursday : [8 8 Pakistan 8] [9 9 Saeed Anwar 9] c [
10 10 Mullally 10] b [11 11 Irani 11] 57 Aamir Sohail b [12 12
Croft 12] 48 [13 13 Ijaz Ahmed 13] c [14 14 Irani 14] b [15 15
Mullally 15] 48 [16 16 Wasim Akram 16] b [17 17 Croft 17] 6 [1
8 18 Inzamam-ul-Haq 18] not out 37 [19 19 Moin Khan 19] b [20
20 Gough 20] 10 [21 21 Salim Malik 21] not out 6 Extras (b-2
lb-4 w-7) 13 Total (for 5 wickets , innings closed) 225
Fall : 1-82 2-141 3-160 4-174 5-203 . Did Not Bat : [22 22 Mu-
shtaq Ahmed 22] , [23 23 Waqar Younis 23] , [24 24 Ata-ur-
Rehman 24] , [25 25 Saqlain Mushtaq 25] . Bowling : [26 26 Go-
ugh 26] 10-0-44-1 , [27 27 Mullally 27] 10-3-31-1 , 1 Headley
10-0-52-0 , [28 28 Irani 28] 10-0-56-1 , [29 29 Croft 29]
10-1-36-2 . [30 30 England 30] N. Knight c [31 31 Moin Khan 31]
] b [32 32 Wasim Akram 32] 26 A. Stewart lbw b [33 33 Waqar Yo-
unis 33] 48 M. Atherton b [34 34 Wasim Akram 34] 65 G. Thorpe
st [35 35 Moin Khan 35] b Aamir Sohail 23 M. Maynard b [36 36
Wasim Akram 36] 41 G. Lloyd not out 2 R. Irani not out 6
Extras (lb-4 w-7 nb-4) 15 Total (for 5 wickets , 46.4
overs) 226 Fall of wickets : 1-57 2-98 3-146 4-200 5-220 .
Did not bat : R. Croft , [37 37 D. Gough 37] , D. Headley ,
A. Mullally . Bowling : [38 38 Wasim Akram 38] 9.4-1-45-3 , [3
9 39 Waqar Younis 39] 7-0-28-1 , [40 40 Saqlain Mushtaq 40]
10-1-54-0 , [41 41 Ata-ur-Rehman 41] 3-0-14-0 , [42 42 Mushtaq
Ahmed 42] 10-0-52-0 , Aamir Sohail 7-1-29-1 . Result : [43 43
England 43] won by five wickets . Second match : August 31 , [
44 44 Edgbaston 44] ([45 45 Birmingham 45]) Third :
September 1 , [46 46 Trent Bridge 46] ([47 47 Nottingham 47]
)
1: 759125 England_cricket_team
2: 759295 Pakistan_national_cricket_team
3: 20206 Manchester
4: 9316 England
5: 759125 England_cricket_team
6: 759295 Pakistan_national_cricket_team
7: 660597 Old_Trafford_Cricket_Ground
8: 759295 Pakistan_national_cricket_team
9: 894437 Saeed_Anwar
10: 2222543 Alan_Mullally
11: 2053953 Ronnie_Irani
12: 1160680 Robert_Croft
13: 938863 Ijaz_Ahmed_(cricketer)
14: 2053953 Ronnie_Irani
15: 2222543 Alan_Mullally
16: 501393 Wasim_Akram

17: 1160680 Robert_Croft
18: 896403 Inzamam-ul-Haq
19: 554132 Moin_Khan
20: 413178 Darren_Gough
21: 938861 Saleem_Malik
22: 896405 Mushtaq_Ahmed
23: 894433 Waqar_Younis
24: 3976751 Ata-ur-Rehman
25: 473185 Saqlain_Mushtaq
26: 413178 Darren_Gough
27: 2222543 Alan_Mullally
28: 2053953 Ronnie_Irani
29: 1160680 Robert_Croft
30: 759125 England_cricket_team
31: 554132 Moin_Khan
32: 501393 Wasim_Akram
33: 894433 Waqar_Younis
34: 501393 Wasim_Akram
35: 554132 Moin_Khan
36: 501393 Wasim_Akram
37: 413178 Darren_Gough
38: 501393 Wasim_Akram
39: 894433 Waqar_Younis
40: 473185 Saqlain_Mushtaq
41: 3976751 Ata-ur-Rehman
42: 896405 Mushtaq_Ahmed
43: 759125 England_cricket_team
44: 2641030 Edgbaston_Cricket_Ground
45: 13882056 Birmingham
46: 849500 Trent_Bridge
47: 39470 Nottingham
1: 672220 Dean_Headley
 [winner: 3903963 Ron_Headley, pem_pos=19,
final_scores=-0.17545041, final_local_scores=-0.20048088,
pem=-3.2580965, similarity=15.824999 |
 RECALL MISS: 672220 Dean_Headley]
1: 759125 England_cricket_team
 [pem_pos=2, final_scores=0.6788818,
final_local_scores=0.648588, pem=-3.3054674,
similarity=41.420517]
2: 759295 Pakistan_national_cricket_team
 [pem_pos=1, final_scores=0.79095584,
final_local_scores=0.7782648, pem=-2.7556365,
similarity=42.300266]
3: 20206 Manchester
 [pem_pos=0, final_scores=0.48805898,
final_local_scores=0.57458556, pem=-0.3209542,
similarity=22.981892]
4: 9316 England
 [pem_pos=0, final_scores=0.52831143,
final_local_scores=0.6214681, pem=-0.19053856,
similarity=23.67174]

```
5: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.7897553,
final_local_scores=0.76970035, pem=-3.3054674,
similarity=45.03475]
6: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.7322035,
final_local_scores=0.71408683, pem=-2.7556365,
similarity=40.385067]
7: 660597 Old_Trafford_Cricket_Ground
    [pem_pos=1, final_scores=0.5623866,
final_local_scores=0.5504525, pem=-0.6437659,
similarity=24.017212]
8: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.8352781,
final_local_scores=0.82668024, pem=-2.7556365,
similarity=43.74508]
9: 894437 Saeed_Anwar
    [pem_pos=0, final_scores=0.8627034,
final_local_scores=0.88340575, pem=0.0, similarity=30.452303]
10: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.8180267,
final_local_scores=0.84791636, pem=-0.73930055,
similarity=33.413662]
11: 2053953 Ronnie_Irani
    [pem_pos=2, final_scores=0.6730327,
final_local_scores=0.6854526, pem=-2.5319417,
similarity=38.314075]
12: 1160680 Robert_Croft
    [pem_pos=7, final_scores=0.93432885,
final_local_scores=0.9737347, pem=-3.16984,
similarity=50.385983]
13: 938863 Ijaz_Ahmed_(cricketer)
    [pem_pos=0, final_scores=0.62160903,
final_local_scores=0.6256857, pem=0.0, similarity=22.761421]
14: 2053953 Ronnie_Irani
    [pem_pos=2, final_scores=0.82350355,
final_local_scores=0.84981906, pem=-2.5319417,
similarity=43.2191]
15: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.7493533,
final_local_scores=0.77290106, pem=-0.73930055,
similarity=31.175053]
16: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.6715457,
final_local_scores=0.65279347, pem=0.0, similarity=23.570372]
17: 1160680 Robert_Croft
    [pem_pos=7, final_scores=0.85146624,
final_local_scores=0.88322, pem=-3.16984, similarity=47.68484]
18: 896403 Inzamam-ul-Haq
    [pem_pos=0, final_scores=1.1129018,
final_local_scores=1.1471701, pem=0.0, similarity=38.323555]
19: 554132 Moin_Khan
```

```
[pem_pos=0, final_scores=0.64089483,
final_local_scores=0.6726843, pem=0.0, similarity=24.163954]
20: 413178 Darren_Gough
    [pem_pos=0, final_scores=1.3379265,
final_local_scores=1.4155908, pem=0.0, similarity=46.333767]
21: 938861 Saleem_Malik
    [pem_pos=0, final_scores=0.76193875,
final_local_scores=0.7930541, pem=0.0, similarity=27.75603]
22: 896405 Mushtaq_Ahmed
    [pem_pos=0, final_scores=0.6710699,
final_local_scores=0.6760732, pem=0.0, similarity=24.265085]
23: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.7410668,
final_local_scores=0.74056655, pem=0.0, similarity=26.189695]
24: 3976751 Ata-ur-Rehman
    [pem_pos=0, final_scores=0.5146905,
final_local_scores=0.5766027, pem=0.0, similarity=21.296688]
25: 473185 Saqlain_Mushtaq
    [pem_pos=0, final_scores=0.66057116,
final_local_scores=0.65490097, pem=0.0, similarity=23.633265]
26: 413178 Darren_Gough
    [pem_pos=0, final_scores=1.3337862,
final_local_scores=1.4110682, pem=0.0, similarity=46.198803]
27: 2222543 Alan_Mullally
    [pem_pos=1, final_scores=0.6562413,
final_local_scores=0.6711903, pem=-0.73930055,
similarity=28.139803]
28: 2053953 Ronnie_Irani
    [pem_pos=2, final_scores=0.6556793,
final_local_scores=0.66649663, pem=-2.5319417,
similarity=37.748394]
29: 1160680 Robert_Croft
    [pem_pos=7, final_scores=0.80659217,
final_local_scores=0.8342018, pem=-3.16984,
similarity=46.222042]
30: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.7179666,
final_local_scores=0.6912822, pem=-3.3054674,
similarity=42.6946]
31: 554132 Moin_Khan
    [pem_pos=0, final_scores=0.6201996,
final_local_scores=0.65007794, pem=0.0, similarity=23.489336]
32: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.7385603,
final_local_scores=0.7259968, pem=0.0, similarity=25.754906]
33: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.7388842,
final_local_scores=0.73818237, pem=0.0, similarity=26.118547]
34: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.6828243,
final_local_scores=0.6651136, pem=0.0, similarity=23.93803]
35: 554132 Moin_Khan
```

```
[pem_pos=0, final_scores=0.5709649,
final_local_scores=0.59629655, pem=0.0, similarity=21.884392]
36: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.68819505,
final_local_scores=0.67098033, pem=0.0, similarity=24.113104]
37: 413178 Darren_Gough
    [pem_pos=0, final_scores=0.69326526,
final_local_scores=0.70766735, pem=0.0, similarity=25.207918]
38: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.78352314,
final_local_scores=0.77511173, pem=0.0, similarity=27.220592]
39: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.7405625,
final_local_scores=0.7400156, pem=0.0, similarity=26.173256]
40: 473185 Saqlain_Mushtaq
    [pem_pos=0, final_scores=0.66281646,
final_local_scores=0.65735364, pem=0.0, similarity=23.706457]
41: 3976751 Ata-ur-Rehman
    [pem_pos=0, final_scores=0.4845177,
final_local_scores=0.5436434, pem=0.0, similarity=20.313116]
42: 896405 Mushtaq_Ahmed
    [pem_pos=0, final_scores=0.5582904,
final_local_scores=0.55287874, pem=0.0, similarity=20.588717]
43: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.77857345,
final_local_scores=0.7574859, pem=-3.3054674,
similarity=44.67025]
44: 2641030 Edgbaston_Cricket_Ground
    [pem_pos=1, final_scores=0.4330067,
final_local_scores=0.40309888, pem=-0.4505704,
similarity=18.56926]
45: 13882056 Birmingham
    [pem_pos=0, final_scores=0.38282752,
final_local_scores=0.4544627, pem=-0.40171203,
similarity=19.836359]
46: 849500 Trent_Bridge
    [pem_pos=0, final_scores=0.59380895,
final_local_scores=0.58345354, pem=0.0, similarity=21.501131]
47: 39470 Nottingham
    [pem_pos=0, final_scores=0.49950832,
final_local_scores=0.56411916, pem=-0.29126778,
similarity=22.508114]
```

Sample: b'807_LOMBARDI'

```
[1 1 LOMBARDI 1] WINS THIRD STAGE OF TOUR OF NETHERLANDS . [2
2 DOETINCHEM 2] , [3 3 Netherlands 3] 1996-08-29 Leading
results and overall standings after the 122 kilometre third
stage of the [4 4 Tour of the Netherlands 4] on Thursday
between [5 5 Almere 5] and [6 6 Doetinchem 6] . 1. [7 7 Giova
nni Lombardi 7] ( [8 8 Italy 8] ) [9 9 Polti 9] 2 hours 35
minutes 29 seconds 2. [10 10 Rolf Sorensen 10] ( [11 11 Denma
rk 11] ) [12 12 Rabobank 12] 3. [13 13 Lance Armstrong 13] ( [
```

14 14 U.S. 14]) [15 15 Motorola 15] 4. [16 16 Maarten den Bakker 16] ([17 17 Netherlands 17]) [18 18 TVM 18] all same time 5. [19 19 Marco Lietti 19] ([20 20 Italy 20]) MG-Technogym 1 second behind 6. Hans de Clerq ([21 21 Belgium 21]) Palmans 27 seconds 7. Marty Jemison ([22 22 U.S. 22]) 1 U.S. Postal 8. [23 23 Servais Knaven 23] ([24 24 Netherlands 24]) [25 25 TVM 25] 9. [26 26 Olaf Ludwig 26] ([27 27 Germany 27]) Telekom all same time 10. [28 28 Jeroen Blijlevens 28] ([29 29 Netherlands 29]) [30 30 TVM 30] 31 Leading overall placings after three stages : 1. Sorensen 10.57:33 2. [31 31 Lombardi 31] 1 second behind 3. [32 32 Armstrong 32] 2 seconds 4. Den Bakker 7 5. [33 33 Lietti 33] 8 6. Federico Colonna ([34 34 Italy 34]) Mapei 27 7. [35 35 Max van Heeswijk 35] ([36 36 Netherlands 36]) [37 37 Motorola 37] 28 8. Sven Teutenberg ([38 38 Germany 38]) 2 U.S. Postal 31 9. [39 39 Johan Capiot 39] ([40 40 Belgium 40]) [41 41 Collstrop 41] 32 10. [42 42 Jans Koerts 42] ([43 43 Netherlands 43]) Palmans 34 The race continues on Thursday afternoon with a 19.6 fourth-stage Doetinchem-Doetinchem time trial

1: 3064824 Giovanni_Lombardi
2: 118600 Doetinchem
3: 21148 Netherlands
4: 11065220 Ronde_van_Nederland
5: 118444 Almere
6: 118600 Doetinchem
7: 3064824 Giovanni_Lombardi
8: 14532 Italy
9: 12164194 Team_Polti
10: 2611279 Rolf_Sørensen
11: 76972 Denmark
12: 2354465 Belkin_Pro_Cycling_Team
13: 23243880 Lance_Armstrong
14: 3434750 United_States
15: 2290911 7-Eleven_(cycling_team)
16: 11100389 Maarten_den_Bakker
17: 21148 Netherlands
18: 2642474 TVM_(cycling_team)
19: 11099892 Marco_Lietti
20: 14532 Italy
21: 3343 Belgium
22: 3434750 United_States
23: 8638361 Servais_Knaven
24: 21148 Netherlands
25: 2642474 TVM_(cycling_team)
26: 2107369 Olaf_Ludwig
27: 11867 Germany
28: 11066258 Jeroen_Blijlevens
29: 21148 Netherlands
30: 2642474 TVM_(cycling_team)
31: 3064824 Giovanni_Lombardi
32: 23243880 Lance_Armstrong
33: 11099892 Marco_Lietti

34: 14532 Italy
35: 6558171 Max_van_Heeswijk
36: 21148 Netherlands
37: 2290911 7-Eleven_(cycling_team)
38: 11867 Germany
39: 13573637 Johan_Capiot
40: 3343 Belgium
41: 6932095 Cycle_Collstrop
42: 17022543 Jans_Koerts
43: 21148 Netherlands

1: 796027 U.S._Postal_Service_Pro_Cycling_Team
[winner: 1148045 Postal_(video_game), pem_pos=0,
final_scores=-1.0360705, final_local_scores=-0.9933476,
pem=-0.8678312, similarity=-20.834358]
RECALL MISS: 796027
U.S._Postal_Service_Pro_Cycling_Team]
[winner: 3434750 United_States, pem_pos=0,
final_scores=-1.1789185, final_local_scores=-1.2924824,
pem=-0.14075686, similarity=-33.715084]
RECALL MISS: 796027 U.S._Postal_Service_Pro_Cycling_Team]
2: 796027 U.S._Postal_Service_Pro_Cycling_Team
[winner: 1148045 Postal_(video_game), pem_pos=0,
final_scores=-1.0264102, final_local_scores=-0.9827952,
pem=-0.8678312, similarity=-20.519453]
RECALL MISS: 796027
U.S._Postal_Service_Pro_Cycling_Team]
[winner: 3268207 United_States_Reports, pem_pos=1,
final_scores=-1.134615, final_local_scores=-1.0610125,
pem=-2.4891596, similarity=-14.036578]
RECALL MISS: 796027 U.S._Postal_Service_Pro_Cycling_Team]
1: 3064824 Giovanni_Lombardi
[pem_pos=0, final_scores=0.4046179,
final_local_scores=0.45978653, pem=0.0, similarity=17.81066]
2: 118600 Doetinchem
[pem_pos=0, final_scores=0.44466448,
final_local_scores=0.49200904, pem=0.0, similarity=18.772243]
3: 21148 Netherlands
[pem_pos=2, final_scores=0.56179667,
final_local_scores=0.5479599, pem=-0.24940795,
similarity=21.79825]
4: 11065220 Ronde_van_Nederland
[pem_pos=0, final_scores=0.56023425,
final_local_scores=0.64160067, pem=0.0, similarity=23.236357]
5: 118444 Almere
[pem_pos=0, final_scores=0.45370215,
final_local_scores=0.49007404, pem=0.0, similarity=18.7145]
6: 118600 Doetinchem
[pem_pos=0, final_scores=0.44079202,
final_local_scores=0.48777896, pem=0.0, similarity=18.64601]
7: 3064824 Giovanni_Lombardi
[pem_pos=0, final_scores=0.5776283,
final_local_scores=0.6487741, pem=0.0, similarity=23.450426]

```
8: 14532 Italy
    [pem_pos=0, final_scores=0.75628656,
final_local_scores=0.803932, pem=-0.20612453,
similarity=29.201586]
9: 12164194 Team Polti
    [pem_pos=0, final_scores=0.46871334,
final_local_scores=0.5496347, pem=0.0, similarity=20.491909]
10: 2611279 Rolf Sørensen
    [pem_pos=0, final_scores=0.7995892,
final_local_scores=0.895215, pem=0.0, similarity=30.804714]
11: 76972 Denmark
    [pem_pos=0, final_scores=0.58413726,
final_local_scores=0.61161613, pem=-0.26762357,
similarity=23.796936]
12: 2354465 Belkin_Pro_Cycling_Team
    [pem_pos=0, final_scores=0.5013235,
final_local_scores=0.5410727, pem=0.0, similarity=20.236403]
13: 23243880 Lance_Armstrong
    [pem_pos=0, final_scores=0.5711763,
final_local_scores=0.6136943, pem=0.0, similarity=22.403576]
14: 3434750 United_States
    [pem_pos=0, final_scores=0.47226995,
final_local_scores=0.51830024, pem=-0.14075686,
similarity=20.322285]
15: 2290911 7-Eleven_(cycling_team)
    [pem_pos=1, final_scores=0.66801614,
final_local_scores=0.7526603, pem=-1.7647687,
similarity=36.147682]
16: 11100389 Maarten_den_Bakker
    [pem_pos=0, final_scores=0.5013258,
final_local_scores=0.6071175, pem=0.0, similarity=22.20731]
17: 21148 Netherlands
    [pem_pos=2, final_scores=0.609626,
final_local_scores=0.6002062, pem=-0.24940795,
similarity=23.357382]
18: 2642474 TVM_(cycling_team)
    [pem_pos=1, final_scores=0.69434357,
final_local_scores=0.8120818, pem=-0.84397006,
similarity=32.913494]
19: 11099892 Marco_Lietti
    [pem_pos=0, final_scores=0.82300013,
final_local_scores=0.9252799, pem=0.0, similarity=31.701912]
20: 14532 Italy
    [pem_pos=0, final_scores=0.7376258,
final_local_scores=0.783548, pem=-0.20612453,
similarity=28.593287]
21: 3343 Belgium
    [pem_pos=0, final_scores=0.73499316,
final_local_scores=0.7377913, pem=-0.24180195,
similarity=27.421835]
22: 3434750 United_States
    [pem_pos=0, final_scores=0.471348,
```

```
final_local_scores=0.51729316, pem=-0.14075686,
similarity=20.292229]
23: 8638361 Servais_Knaven
    [pem_pos=0, final_scores=0.5627479,
final_local_scores=0.67602926, pem=0.0, similarity=24.263775]
24: 21148 Netherlands
    [pem_pos=2, final_scores=0.6195003,
final_local_scores=0.6109924, pem=-0.24940795,
similarity=23.679264]
25: 2642474 TVM_(cycling_team)
    [pem_pos=1, final_scores=0.7418314,
final_local_scores=0.8639551, pem=-0.84397006,
similarity=34.461494]
26: 2107369 Olaf_Ludwig
    [pem_pos=0, final_scores=0.6241628,
final_local_scores=0.7185967, pem=0.0, similarity=25.534071]
27: 11867 Germany
    [pem_pos=0, final_scores=0.868494,
final_local_scores=0.91436654, pem=-0.15529114,
similarity=32.220734]
28: 11066258 Jeroen_Blijlevens
    [pem_pos=0, final_scores=0.7675907,
final_local_scores=0.86211276, pem=0.0, similarity=29.816877]
29: 21148 Netherlands
    [pem_pos=2, final_scores=0.63629466,
final_local_scores=0.62933767, pem=-0.24940795,
similarity=24.226725]
30: 2642474 TVM_(cycling_team)
    [pem_pos=1, final_scores=0.63119715,
final_local_scores=0.743104, pem=-0.84397006,
similarity=30.855057]
31: 3064824 Giovanni_Lombardi
    [pem_pos=0, final_scores=0.54783374,
final_local_scores=0.6162281, pem=0.0, similarity=22.479189]
32: 23243880 Lance_Armstrong
    [pem_pos=0, final_scores=1.0254543,
final_local_scores=1.1169465, pem=0.0, similarity=37.421623]
33: 11099892 Marco_Lietti
    [pem_pos=0, final_scores=0.47268575,
final_local_scores=0.54261464, pem=0.0, similarity=20.282415]
34: 14532 Italy
    [pem_pos=0, final_scores=0.773248,
final_local_scores=0.8224599, pem=-0.20612453,
similarity=29.754494]
35: 6558171 Max_van_Heeswijk
    [pem_pos=0, final_scores=0.5853073,
final_local_scores=0.69051194, pem=0.0, similarity=24.695967]
36: 21148 Netherlands
    [pem_pos=2, final_scores=0.6518631,
final_local_scores=0.6463438, pem=-0.24940795,
similarity=24.734222]
37: 2290911 7-Eleven_(cycling_team)
```

```
[pem_pos=1, final_scores=0.75558287,
final_local_scores=0.84831357, pem=-1.7647687,
similarity=39.002167]
38: 11867 Germany
    [pem_pos=0, final_scores=0.83015317,
final_local_scores=0.8724851, pem=-0.15529114,
similarity=30.970907]
39: 13573637 Johan_Capiot
    [pem_pos=0, final_scores=0.60893154,
final_local_scores=0.75130886, pem=0.0, similarity=26.510267]
40: 3343 Belgium
    [pem_pos=0, final_scores=0.73863536,
final_local_scores=0.7417699, pem=-0.24180195,
similarity=27.540564]
41: 6932095 Cycle_Collstrop
    [pem_pos=0, final_scores=0.5399875,
final_local_scores=0.64004076, pem=0.0, similarity=23.189806]
42: 17022543 Jans_Koerts
    [pem_pos=0, final_scores=0.6132138,
final_local_scores=0.73498607, pem=0.0, similarity=26.023163]
43: 21148 Netherlands
    [pem_pos=2, final_scores=0.6304335,
final_local_scores=0.62293524, pem=-0.24940795,
similarity=24.035662]
```

Sample: b'808_CRICKET'

CRICKET - [1 1 ENGLAND 1] V [2 2 PAKISTAN 2] TOSS AND
TEAMS . [3 3 MANCHESTER 3] , [4 4 England 4] 1996-08-29 [5 5 P
akistan 5] won the toss and elected to bat in the first one-
day cricket international between [6 6 England 6] and [7 7 Pa
kistan 7] at [8 8 Old Trafford 8] on Thursday . Teams : [9 9 E
ngland 9] - [10 10 Mike Atherton 10] (captain) , [11 11 Nick
Knight 11] , [12 12 Alec Stewart 12] , [13 13 Graham Thorpe 13
] , [14 14 Matthew Maynard 14] , [15 15 Graham Lloyd 15] , [16
16 Ronnie Irani 16] , [17 17 Robert Croft 17] , [18 18 Darren
Gough 18] , [19 19 Dean Headley 19] , [20 20 Alan Mullally 20]
. [21 21 Pakistan 21] : Aamir Sohail , [22 22 Saeed Anwar 22]
, [23 23 Ijaz Ahmed 23] , [24 24 Inzamam-ul-Haq 24] , [25 25 S
alim Malik 25] , [26 26 Wasim Akram 26] (captain) , [27 27 M
oin Khan 27] , [28 28 Mushtaq Ahmed 28] , [29 29 Waqar Younis
29] , [30 30 Ata-ur-Rehman 30] , [31 31 Saqlain Mushtaq 31] .

```
1: 759125 England_cricket_team
2: 759295 Pakistan_national_cricket_team
3: 20206 Manchester
4: 9316 England
5: 759295 Pakistan_national_cricket_team
6: 759125 England_cricket_team
7: 759295 Pakistan_national_cricket_team
8: 660597 Old_Trafford_Cricket_Ground
9: 759125 England_cricket_team
10: 159024 Michael_Atherton
11: 994718 Nick_Knight
```

```
12: 896414 Alec_Stewart
13: 397259 Graham_Thorpe
14: 2512623 Matthew_Maynard
15: 3750666 Graham_Lloyd
16: 2053953 Ronnie_Irani
17: 1160680 Robert_Croft
18: 413178 Darren_Gough
19: 672220 Dean_Headley
20: 2222543 Alan_Mullally
21: 759295 Pakistan_national_cricket_team
22: 894437 Saeed_Anwar
23: 938863 Ijaz_Ahmed_(cricketer)
24: 896403 Inzamam-ul-Haq
25: 938861 Saleem_Malik
26: 501393 Wasim_Akram
27: 554132 Moin_Khan
28: 896405 Mushtaq_Ahmed
29: 894433 Waqar_Younis
30: 3976751 Ata-ur-Rehman
31: 473185 Saqlain_Mushtaq
1: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.6492692,
final_local_scores=0.60964245, pem=-3.3054674,
similarity=40.258305]
2: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.78078395,
final_local_scores=0.7619048, pem=-2.7556365,
similarity=41.81205]
3: 20206 Manchester
    [pem_pos=0, final_scores=0.47169942,
final_local_scores=0.5570234, pem=-0.3209542,
similarity=22.457802]
4: 9316 England
    [pem_pos=0, final_scores=0.4698364,
final_local_scores=0.55416566, pem=-0.19053856,
similarity=21.6633]
5: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.8255082,
final_local_scores=0.81075925, pem=-2.7556365,
similarity=43.269966]
6: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.81047016,
final_local_scores=0.78573006, pem=-3.3054674,
similarity=45.51311]
7: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.84852475,
final_local_scores=0.8359013, pem=-2.7556365,
similarity=44.020256]
8: 660597 Old_Trafford_Cricket_Ground
    [pem_pos=1, final_scores=0.64253074,
final_local_scores=0.6378501, pem=-0.6437659,
similarity=26.625332]
```

```
9: 759125 England_cricket_team
    [pem_pos=2, final_scores=0.73826903,
final_local_scores=0.7068612, pem=-3.3054674,
similarity=43.159508]
10: 159024 Michael_Atherton
    [pem_pos=0, final_scores=0.9763306,
final_local_scores=1.0047145, pem=0.0, similarity=34.072395]
11: 994718 Nick_Knight
    [pem_pos=1, final_scores=0.6810507,
final_local_scores=0.7391036, pem=-0.55008924,
similarity=29.137508]
12: 896414 Alec_Stewart
    [pem_pos=0, final_scores=0.9088388,
final_local_scores=0.93582195, pem=0.0, similarity=32.016506]
13: 397259 Graham_Thorpe
    [pem_pos=0, final_scores=0.7324342,
final_local_scores=0.74508405, pem=0.0, similarity=26.324507]
14: 2512623 Matthew_Maynard
    [pem_pos=0, final_scores=0.67337495,
final_local_scores=0.70612293, pem=0.0, similarity=25.161829]
15: 3750666 Graham_Lloyd
    [pem_pos=0, final_scores=0.526015,
final_local_scores=0.5949214, pem=0.0, similarity=21.843355]
16: 2053953 Ronnie_Irani
    [pem_pos=0, final_scores=0.6573513,
final_local_scores=0.67845017, pem=0.0, similarity=24.33602]
17: 1160680 Robert_Croft
    [pem_pos=0, final_scores=1.0441198,
final_local_scores=1.1031995, pem=0.0, similarity=37.011387]
18: 413178 Darren_Gough
    [pem_pos=0, final_scores=0.91872567,
final_local_scores=0.966912, pem=0.0, similarity=32.944294]
19: 672220 Dean_Headley
    [pem_pos=0, final_scores=0.5048967,
final_local_scores=0.5688712, pem=0.0, similarity=21.065964]
20: 2222543 Alan_Mullally
    [pem_pos=0, final_scores=0.76012665,
final_local_scores=0.7911474, pem=0.0, similarity=27.69913]
21: 759295 Pakistan_national_cricket_team
    [pem_pos=1, final_scores=0.8026261,
final_local_scores=0.78576404, pem=-2.7556365,
similarity=42.52406]
22: 894437 Saeed_Anwar
    [pem_pos=0, final_scores=0.8330818,
final_local_scores=0.84875524, pem=0.0, similarity=29.418262]
23: 938863 Ijaz_Ahmed_(cricketer)
    [pem_pos=0, final_scores=0.6778937,
final_local_scores=0.6874791, pem=0.0, similarity=24.60546]
24: 896403 Inzamam-ul-Haq
    [pem_pos=0, final_scores=0.89982766,
final_local_scores=0.9152424, pem=0.0, similarity=31.402372]
25: 938861 Saleem_Malik
```

```
[pem_pos=0, final_scores=0.77297884,
final_local_scores=0.80391955, pem=0.0, similarity=28.080276]
26: 501393 Wasim_Akram
    [pem_pos=0, final_scores=0.81978136,
final_local_scores=0.8236972, pem=0.0, similarity=28.67048]
27: 554132 Moin_Khan
    [pem_pos=0, final_scores=0.8045792,
final_local_scores=0.8600841, pem=0.0, similarity=29.756338]
28: 896405 Mushtaq_Ahmed
    [pem_pos=0, final_scores=0.64247704,
final_local_scores=0.6504905, pem=0.0, similarity=23.501648]
29: 894433 Waqar_Younis
    [pem_pos=0, final_scores=0.7687059,
final_local_scores=0.77561706, pem=0.0, similarity=27.235674]
30: 3976751 Ata-ur-Rehman
    [pem_pos=0, final_scores=0.5509667,
final_local_scores=0.6247842, pem=0.0, similarity=22.734518]
31: 473185 Saqlain_Mushtaq
    [pem_pos=0, final_scores=0.698509,
final_local_scores=0.6990545, pem=0.0, similarity=24.950893]
```

Sample: b'809_CRICKET'

CRICKET - [1 1 ENGLISH COUNTY CHAMPIONSHIP 1] SCORES . [2 2 L
ONDON 2] 1996-08-29 Close of play scores in [3 3 English 3]
county championship matches on Thursday : [4 4 Tunbridge Wells
4] : [5 5 Nottinghamshire 5] 40-3 v [6 6 Kent 6] [7 7 London 7
] ([8 8 The Oval 8]) : [9 9 Warwickshire 9] 195 (A. Giles
50 ; B. Julian 4-66 , C. Lewis 4-45) , [10 10 Surrey 10]
82-0 . [11 11 Hove 11] : [12 12 Sussex 12] 285-6 (W. Athey
111) v Lancashire Leeds ([13 13 Headingley 13]) : [14 14 Yo
rkshire 14] 290 (C. White 76 , M. Moxon 59 , R. Blakey
57) , [15 15 Essex 15] 79-2 . [16 16 Chester-le-Street 16] :
[17 17 Glamorgan 17] 259 (P. Cottney 81 ; M. Saggers 6-65)
and 8-0 , [18 18 Durham 18] 114 (S. Watkin 4-28) 1 Chesterf
ield : [19 19 Worcestershire 19] 238 (W. Weston 100 not
out , V. Solanki 58 ; A. Harris 4-31) , [20 20 Derbyshire 20]
166-1 (K. Barnett 83 not out) [21 21 Portsmouth 21] : [22 22
Middlesex 22] 199 and 226-1 (J. Pooley 106 not out , M.
Ramprakash 81 not out) , [23 23 Hampshire 23] 232 (A.
Fraser 5-55 , R. Fay 4-77) [24 24 Leicester 24] : [25 25 Som
erset 25] 83 (D. Millns 4-35) , [26 26 Leicestershire 26]
202-5 2 Bristol : [27 27 Gloucestershire 27] 183 (J. Russell
50) , [28 28 Northamptonshire 28] 123-4 (K. Curran 51 not
out) .

1: 683599 County_Championship
2: 17867 London
3: 9316 England
4: 10877349 Nevill_Ground
5: 667341 Nottinghamshire_County_Cricket_Club
6: 1177719 Kent_County_Cricket_Club
7: 17867 London
8: 387687 The_Oval

9: 667292 Warwickshire_County_Cricket_Club
10: 766575 Surrey_County_Cricket_Club
11: 2642672 County_Cricket_Ground,_Hove
12: 1763691 Sussex_County_Cricket_Club
13: 1876060 Headingley_Stadium
14: 869945 Yorkshire_County_Cricket_Club
15: 1622252 Essex_County_Cricket_Club
16: 2311656 Riverside_Ground
17: 1622320 Glamorgan_County_Cricket_Club
18: 1176371 Durham_County_Cricket_Club
19: 1156428 Worcestershire_County_Cricket_Club
20: 1829984 Derbyshire_County_Cricket_Club
21: 24207 Portsmouth
22: 822103 Middlesex_County_Cricket_Club
23: 1919370 Hampshire_County_Cricket_Club
24: 51155 Leicester
25: 1622178 Somerset_County_Cricket_Club
26: 1622318 Leicestershire_County_Cricket_Club
27: 1622322 Gloucestershire_County_Cricket_Club
28: 530308 Northamptonshire_County_Cricket_Club
1: 10756265 Queen's_Park,_Chesterfield
[winner: 451126 Chesterfield_F.C., pem_pos=1,
final_scores=-0.11984621, final_local_scores=-0.14740051,
pem=-1.2678006, similarity=6.585484]
RECALL MISS: 10756265 Queen's_Park,_Chesterfield]
2: 2642664 County_Cricket_Ground,_Bristol
[winner: 36741 Bristol, pem_pos=0,
final_scores=-0.14902784, final_local_scores=-0.19529194,
pem=-0.43489033, similarity=0.6268134]
RECALL MISS: 2642664 County_Cricket_Ground,_Bristol]
1: 683599 County_Championship
[pem_pos=0, final_scores=0.36427766,
final_local_scores=0.3070232, pem=0.0, similarity=13.251898]
2: 17867 London
[pem_pos=0, final_scores=0.6763558,
final_local_scores=0.744485, pem=-0.09086737,
similarity=26.800781]
3: 9316 England
[pem_pos=1, final_scores=0.5332629,
final_local_scores=0.5621906, pem=-1.3527626,
similarity=28.223133]
4: 10877349 Nevill_Ground
[pem_pos=8, final_scores=0.3312449,
final_local_scores=0.37397403, pem=-2.2777648,
similarity=27.636683]
5: 667341 Nottinghamshire_County_Cricket_Club
[pem_pos=1, final_scores=0.59132665,
final_local_scores=0.53722763, pem=-1.1067253,
similarity=26.140202]
6: 1177719 Kent_County_Cricket_Club
[pem_pos=1, final_scores=0.6350917,
final_local_scores=0.58481187, pem=-2.5096242,

```
similarity=35.189392]
7: 17867 London
    [pem_pos=0, final_scores=0.5988376,
final_local_scores=0.65980816, pem=-0.09086737,
similarity=24.273855]
8: 387687 The_Oval
    [pem_pos=0, final_scores=0.58149356,
final_local_scores=0.5481144, pem=0.0, similarity=20.44654]
9: 667292 Warwickshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.7611341,
final_local_scores=0.72294855, pem=-1.1093602,
similarity=31.696812]
10: 766575 Surrey_County_Cricket_Club
    [pem_pos=1, final_scores=0.7739982,
final_local_scores=0.734229, pem=-1.7952642,
similarity=35.763496]
11: 2642672 County_Cricket_Ground,_Hove
    [pem_pos=2, final_scores=0.31427163,
final_local_scores=0.27943248, pem=-2.7078161,
similarity=27.15406]
12: 1763691 Sussex_County_Cricket_Club
    [pem_pos=1, final_scores=0.62050194,
final_local_scores=0.5844664, pem=-1.5566913,
similarity=29.996885]
13: 1876060 Headingley_Stadium
    [pem_pos=0, final_scores=0.39209622,
final_local_scores=0.3809604, pem=-0.30334172,
similarity=17.107952]
14: 869945 Yorkshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6906629,
final_local_scores=0.6439912, pem=-1.7779728,
similarity=32.976585]
15: 1622252 Essex_County_Cricket_Club
    [pem_pos=1, final_scores=0.94992274,
final_local_scores=0.9310561, pem=-2.3459642,
similarity=44.632]
16: 2311656 Riverside_Ground
    [pem_pos=3, final_scores=0.43789297,
final_local_scores=0.44988316, pem=-1.4520724,
similarity=25.411718]
17: 1622320 Glamorgan_County_Cricket_Club
    [pem_pos=1, final_scores=0.8044997,
final_local_scores=0.77394336, pem=-0.7903395,
similarity=31.483715]
18: 1176371 Durham_County_Cricket_Club
    [pem_pos=2, final_scores=0.8653551,
final_local_scores=0.85284597, pem=-2.3952909,
similarity=42.566303]
19: 1156428 Worcestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.66467077,
final_local_scores=0.6191865, pem=-0.9164197,
similarity=27.551105]
```

```
20: 1829984 Derbyshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.7129019,
final_local_scores=0.6713179, pem=-1.4828352,
similarity=32.187065]
21: 24207 Portsmouth
    [pem_pos=0, final_scores=0.39904392,
final_local_scores=0.4414255, pem=-0.75559247,
similarity=21.37176]
22: 822103 Middlesex_County_Cricket_Club
    [pem_pos=1, final_scores=0.75634843,
final_local_scores=0.71417975, pem=-1.5695794,
similarity=33.937878]
23: 1919370 Hampshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.766488,
final_local_scores=0.73310953, pem=-1.9232318,
similarity=36.425995]
24: 51155 Leicester
    [pem_pos=0, final_scores=0.48868102,
final_local_scores=0.4942522, pem=-0.2625849,
similarity=20.267162]
25: 1622178 Somerset_County_Cricket_Club
    [pem_pos=1, final_scores=0.693108,
final_local_scores=0.6484499, pem=-2.4139037,
similarity=36.567936]
26: 1622318 Leicestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.6878347,
final_local_scores=0.64855725, pem=-1.0465667,
similarity=29.135347]
27: 1622322 Gloucestershire_County_Cricket_Club
    [pem_pos=1, final_scores=0.5954592,
final_local_scores=0.5415291, pem=-1.379793,
similarity=27.753551]
28: 530308 Northamptonshire_County_Cricket_Club
    [pem_pos=1, final_scores=0.41225195,
final_local_scores=0.34812737, pem=-1.2069724,
similarity=21.04223]
```

Sample: b'810_RUGBY'

```
[1 1 RUGBY UNION 1] - [2 2 TELFER 2] CONFIRMED FOR [1 1 LIONS
1] COACHING ROLE . [3 3 LONDON 3] 1996-08-30 [4 4 Scotland 4]
's [5 5 Jim Telfer 5] was officially confirmed on Thursday as
assistant coach for the [2 2 British Lions 2] tour to [6 6 So
uth Africa 6] next year . [7 7 Telfer 7] , who has put on
hold his role as [8 8 Scotland 8] team manager for a year ,
will act as assistant to [9 9 Ian McGeechan 9] . The pair
last worked together when [10 10 Scotland 10] won the [11 11 F
ive Nations 11] grand slam in 1990 . The tour party will be
announced towards the end of March .
```

```
1: 25405 Rugby_union
2: 1218828 Jim_Telfer
3: 17867 London
4: 758049 Scotland_national_rugby_union_team
```

```
5: 1218828 Jim_Telfer
6: 17416221 South_Africa
7: 1218828 Jim_Telfer
8: 758049 Scotland_national_rugby_union_team
9: 1211786 Ian_McGeechan
10: 758049 Scotland_national_rugby_union_team
11: 26547 Six_Nations_Championship
1: 951941 Great_Britain_national_rugby_league_team
   [winner: 3914 British_and_Irish_Lions, pem_pos=3,
final_scores=0.24731301, final_local_scores=0.19081348,
pem=-2.2812183, similarity=22.189589]
   gt: 951941 Great_Britain_national_rugby_league_team,
pem_pos=10, final_scores=0.23490982,
final_local_scores=0.2113257, pem=-3.8479764,
similarity=31.321987 ]
2: 951941 Great_Britain_national_rugby_league_team
   [winner: 3914 British_and_Irish_Lions, pem_pos=0,
final_scores=0.21572779, final_local_scores=0.15631142,
pem=-0.11496338, similarity=9.379546]
   gt: 951941 Great_Britain_national_rugby_league_team,
pem_pos=2, final_scores=0.19670738,
final_local_scores=0.16959536, pem=-1.5341516,
similarity=17.49373 ]
1: 25405 Rugby_union
   [pem_pos=0, final_scores=0.46050984,
final_local_scores=0.46308511, pem=0.0, similarity=17.909096]
2: 1218828 Jim_Telfer
   [pem_pos=0, final_scores=0.6014584,
final_local_scores=0.6181505, pem=0.0, similarity=22.536556]
3: 17867 London
   [pem_pos=0, final_scores=0.51241785,
final_local_scores=0.6034232, pem=-0.09086737,
similarity=22.591215]
4: 758049 Scotland_national_rugby_union_team
   [pem_pos=2, final_scores=0.47042412,
final_local_scores=0.4259749, pem=-3.3931751,
similarity=35.25427]
5: 1218828 Jim_Telfer
   [pem_pos=0, final_scores=0.7450864,
final_local_scores=0.7633742, pem=0.0, similarity=26.870321]
6: 17416221 South_Africa
   [pem_pos=0, final_scores=0.37657958,
final_local_scores=0.45716077, pem=-0.26251382,
similarity=19.159893]
7: 1218828 Jim_Telfer
   [pem_pos=0, final_scores=0.7103148,
final_local_scores=0.7370596, pem=0.0, similarity=26.085041]
8: 758049 Scotland_national_rugby_union_team
   [pem_pos=2, final_scores=0.4165191,
final_local_scores=0.36709183, pem=-3.3931751,
similarity=33.497078]
9: 1211786 Ian_McGeechan
```

```
[pem_pos=0, final_scores=0.48719102,
final_local_scores=0.4982028, pem=0.0, similarity=18.957079]
10: 758049 Scotland_national_rugby_union_team
    [pem_pos=2, final_scores=0.4994846,
final_local_scores=0.45771903, pem=-3.3931751,
similarity=36.201576]
11: 26547 Six_Nations_Championship
    [pem_pos=0, final_scores=0.33022898,
final_local_scores=0.3105669, pem=-0.38417557,
similarity=15.446856]
1: 3914 British_and_Irish_Lions
    [pem_pos=3, final_scores=0.24731301,
final_local_scores=0.19081348, pem=-2.2812183,
similarity=22.189589]
2: 3914 British_and_Irish_Lions
    [pem_pos=0, final_scores=0.21572779,
final_local_scores=0.15631142, pem=-0.11496338,
similarity=9.379546]
```

Sample: b'811 ATHLETICS'

ATHLETICS - [1 1 CHRISTIE 1] TO RUN IN [2 2 BERLIN 2] . [3 3 LONDON 3] 1996-08-29 [4 4 Linford Christie 4] has confirmed he will run in a " Dream Team " sprint relay at the [5 5 Berlin 5] grand prix athletics meeting on Friday . A spokeswoman for [6 6 Christie 6] said the former 1 Olympic 100 metres champion had agreed to captain a quality quartet which also includes [7 7 Canada 7] 's [8 8 Donovan Bailey 8] , the current 2 Olympic champion and world record holder , and Namibian [9 9 Frankie Fredericks 9] . [10 10 Christie 10] is retiring from international competition at the end of the season , but [11 11 Berlin 11] promoter Rudi Thiel has persuaded him to join what is intended to be a special tribute to [12 12 Jesse Owens 12] , who won four gold medals 60 years ago at the [13 13 Berlin 13] 3 Olympics .

```
1: 67118 Linford_Christie
2: 3354 Berlin
3: 17867 London
4: 67118 Linford_Christie
5: 3354 Berlin
6: 67118 Linford_Christie
7: 5042916 Canada
8: 67116 Donovan_Bailey
9: 71410 Frankie_Fredericks
10: 67118 Linford_Christie
11: 3354 Berlin
12: 46827 Jesse_Owens
13: 3354 Berlin
1: 82755 1992_Summer_Olympics
    [winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.33997726, final_local_scores=-0.39384952,
pem=-3.2532303, similarity=10.02803|
RECALL MISS: 82755 1992_Summer_Olympics]
```

```
2: 72311 1996_Summer_Olympics
    [winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.33212018, final_local_scores=-0.3852668,
pem=-3.2532303, similarity=10.284156]
    RECALL MISS: 72311 1996_Summer_Olympics]
3: 39721 1936_Summer_Olympics
    [winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.34443653, final_local_scores=-0.39872056,
pem=-2.7333555, similarity=7.055509]
    RECALL MISS: 39721 1936_Summer_Olympics]
1: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.4057972,
final_local_scores=1.465369, pem=0.0, similarity=47.81925]
2: 3354 Berlin
    [pem_pos=0, final_scores=0.29266363,
final_local_scores=0.30256546, pem=-0.15483057,
similarity=13.960862]
3: 17867 London
    [pem_pos=0, final_scores=0.7494367,
final_local_scores=0.78600675, pem=-0.09086737,
similarity=28.039873]
4: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.0830557,
final_local_scores=1.1041529, pem=0.0, similarity=37.039837]
5: 3354 Berlin
    [pem_pos=0, final_scores=0.5949039,
final_local_scores=0.63271695, pem=-0.15483057,
similarity=23.813242]
6: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.6142563,
final_local_scores=1.6930788, pem=0.0, similarity=54.61456]
7: 5042916 Canada
    [pem_pos=0, final_scores=0.7550964,
final_local_scores=0.8492502, pem=-0.11662793,
similarity=30.067274]
8: 67116 Donovan_Bailey
    [pem_pos=0, final_scores=0.48469782,
final_local_scores=0.54602945, pem=0.0, similarity=20.384321]
9: 71410 Frankie_Fredericks
    [pem_pos=0, final_scores=0.7602357,
final_local_scores=0.862828, pem=0.0, similarity=29.838223]
10: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.8155637,
final_local_scores=1.9129767, pem=0.0, similarity=61.176754]
11: 3354 Berlin
    [pem_pos=0, final_scores=0.59296054,
final_local_scores=0.6305942, pem=-0.15483057,
similarity=23.749895]
12: 46827 Jesse_Owens
    [pem_pos=0, final_scores=0.65489954,
final_local_scores=0.71121603, pem=0.0, similarity=25.313818]
13: 3354 Berlin
```

```
[pem_pos=0, final_scores=0.4716696,  
final_local_scores=0.4981023, pem=-0.15483057,  
similarity=19.796072]
```

Sample: b'812_SOCCER'
SOCCER - [1 1 GROBBELAAR 1] NAMED TEMPORARY [2 2 ZIMBABWE 2]
COACH . [3 3 HARARE 3] 1996-08-29 England-based goalkeeper [4
4 Bruce Grobbelaar 4] has been appointed temporary head coach
of the [5 5 Zimbabwe 5] national soccer team for two
international matches , the Zimbabwe Football Association
(ZIFA) said on Thursday . ZIFA vice-chairman Vincent Pamire
said [6 6 Grobbelaar 6] would take charge for a match against
[7 7 Tanzania 7] in [8 8 Harare 8] on September 29 in the
five-nation Castle Cup of Africa tournament and an African
Nations ' Cup first round qualifier against [9 9 Sudan 9] in [
10 10 Khartoum 10] on October 5 . [11 11 Grobbelaar 11] takes
over until a permanent replacement is appointed for [12 12 Zi
mbabwe 12] 's previous coach , [13 13 Switzerland 13] 's Marc
Duvillard . [14 14 Grobbelaar 14] now plays for [15 15 English
15] second division leaders [16 16 Plymouth Argyle 16] after
years in the top flight with [17 17 Liverpool 17] and [18 18 S
outhampton 18] . [19 19 Grobbelaar 19] , fellow goalkeeper [20
20 Hans Segers 20] , retired striker [21 21 John Fashanu 21]
and [22 22 Malaysian 22] businessman Heng Suan Lim pleaded
not guilty in May to charges of giving or accepting bribes to
fix [23 23 English 23] premier league matches . They are due
to stand trial next January .

```
1: 663785 Bruce_Grobbelaar  
2: 1083858 Zimbabwe_national_football_team  
3: 56680 Harare  
4: 663785 Bruce_Grobbelaar  
5: 34399 Zimbabwe  
6: 663785 Bruce_Grobbelaar  
7: 1202658 Tanzania_national_football_team  
8: 56680 Harare  
9: 1095575 Sudan_national_football_team  
10: 17320 Khartoum  
11: 663785 Bruce_Grobbelaar  
12: 1083858 Zimbabwe_national_football_team  
13: 26748 Switzerland  
14: 663785 Bruce_Grobbelaar  
15: 9316 England  
16: 23857 Plymouth_Argyle_F.C.  
17: 18119 Liverpool_F.C.  
18: 81576 Southampton_F.C.  
19: 663785 Bruce_Grobbelaar  
20: 3588980 Hans_Segers  
21: 3830902 John_Fashanu  
22: 3607937 Malaysia  
23: 9316 England  
1: 663785 Bruce_Grobbelaar  
[pem_pos=0, final_scores=0.9244843,
```

```
final_local_scores=0.95149225, pem=0.0, similarity=32.48414]
2: 1083858 Zimbabwe_national_football_team
    [pem_pos=2, final_scores=0.48527664,
final_local_scores=0.48150766, pem=-2.8678126,
similarity=34.054474]
3: 56680 Harare
    [pem_pos=0, final_scores=0.4175294,
final_local_scores=0.44505835, pem=0.0, similarity=17.371141]
4: 663785 Bruce_Grobelaar
    [pem_pos=0, final_scores=0.7842248,
final_local_scores=0.79828, pem=0.0, similarity=27.91198]
5: 34399 Zimbabwe
    [pem_pos=0, final_scores=0.44463938,
final_local_scores=0.49607766, pem=-0.36737105,
similarity=20.891481]
6: 663785 Bruce_Grobelaar
    [pem_pos=0, final_scores=0.5411133,
final_local_scores=0.5327177, pem=0.0, similarity=19.987072]
7: 1202658 Tanzania_national_football_team
    [pem_pos=1, final_scores=0.61935693,
final_local_scores=0.6734809, pem=-2.2414286,
similarity=36.376965]
8: 56680 Harare
    [pem_pos=0, final_scores=0.45635527,
final_local_scores=0.48746973, pem=0.0, similarity=18.636782]
9: 1095575 Sudan_national_football_team
    [pem_pos=1, final_scores=0.5171055,
final_local_scores=0.56690955, pem=-2.8338492,
similarity=36.41834]
10: 17320 Khartoum
    [pem_pos=0, final_scores=0.3807029,
final_local_scores=0.45472342, pem=-0.03966028,
similarity=17.875246]
11: 663785 Bruce_Grobelaar
    [pem_pos=0, final_scores=0.5209169,
final_local_scores=0.5106562, pem=0.0, similarity=19.328712]
12: 1083858 Zimbabwe_national_football_team
    [pem_pos=2, final_scores=0.514913,
final_local_scores=0.51388085, pem=-2.8678126,
similarity=35.020554]
13: 26748 Switzerland
    [pem_pos=0, final_scores=0.46047348,
final_local_scores=0.5280323, pem=-0.18598008,
similarity=20.858639]
14: 663785 Bruce_Grobelaar
    [pem_pos=0, final_scores=0.56353563,
final_local_scores=0.5572107, pem=0.0, similarity=20.71799]
15: 9316 England
    [pem_pos=1, final_scores=0.6277745,
final_local_scores=0.67481476, pem=-1.3527626,
similarity=31.584064]
16: 23857 Plymouth_Argyle_F.C.
```

```
[pem_pos=0, final_scores=0.47319514,
final_local_scores=0.4583187, pem=0.0, similarity=17.766857]
17: 18119 Liverpool_F.C.
    [pem_pos=1, final_scores=0.65287733,
final_local_scores=0.64198756, pem=-1.1702383,
similarity=29.611841]
18: 81576 Southampton_F.C.
    [pem_pos=1, final_scores=0.6604692,
final_local_scores=0.6574636, pem=-1.1503892,
similarity=29.965733]
19: 663785 Bruce_Grobelaar
    [pem_pos=0, final_scores=0.7278662,
final_local_scores=0.7367168, pem=0.0, similarity=26.074812]
20: 3588980 Hans_Segers
    [pem_pos=0, final_scores=0.76316184,
final_local_scores=0.85623944, pem=0.0, similarity=29.641605]
21: 3830902 John_Fashanu
    [pem_pos=0, final_scores=0.7282203,
final_local_scores=0.77558744, pem=0.0, similarity=27.234789]
22: 3607937 Malaysia
    [pem_pos=0, final_scores=0.5229425,
final_local_scores=0.617339, pem=-0.30526873,
similarity=24.17244]
23: 9316 England
    [pem_pos=1, final_scores=0.59626186,
final_local_scores=0.640392, pem=-1.3527626,
similarity=30.556822]
```

Sample: b'813_BASKETBALL'

BASKETBALL - [1 1 OLYMPIAKOS 1] BEAT [2 2 DINAMO 2] 69-60 . [3 3 BELGRADE 3] 1996-08-29 Olympiakos of [4 4 Greece 4] beat [5 5 Russia 5] 's [6 6 Dinamo 6] 69-60 (halftime 35-23) in the third match of an international basketball tournament on Thursday , qualifying for the finals . [7 7 Partizan 7] and [8 8 Red Star 8] of [9 9 Yugoslavia 9] , Alba of [10 10 Germany 1 0] , and [11 11 Benetton 11] of [12 12 Italy 12] are also taking part in the event which continues until Saturday . Add results [13 13 Partizan 13] beat [14 14 Benetton 14] 97-94 (halftime 39-32) . Final [15 15 Partizan 15] v Olympiakos .

```
1: 1229550 Olympiacos_F.C.
2: 968856 FC_Dynamo_Moscow
3: 55904 Belgrade
4: 12108 Greece
5: 25391 Russia
6: 15820894 Dynamo_Sports_Club
7: 5551651 KK_Partizan
8: 5604103 KK_Crvena_zvezda
9: 34244 Yugoslavia
10: 11867 Germany
11: 3450530 Treviso_Basket
12: 14532 Italy
13: 5551651 KK_Partizan
```

```
14: 3450530 Treviso_Basket
15: 5551651 KK_Partizan
1: 1229550 Olympiacos_F.C.
    [pem_pos=0, final_scores=0.22439276,
final_local_scores=0.21244681, pem=-0.4098014,
similarity=12.6581135]
2: 968856 FC_Dynamo_Moscow
    [pem_pos=4, final_scores=0.044434056,
final_local_scores=0.019860774, pem=-2.1006439,
similarity=16.106026]
3: 55904 Belgrade
    [pem_pos=0, final_scores=0.47234184,
final_local_scores=0.46383858, pem=-0.05193398,
similarity=18.214006]
4: 12108 Greece
    [pem_pos=0, final_scores=0.76285857,
final_local_scores=0.81361914, pem=-0.27607638,
similarity=29.871077]
5: 25391 Russia
    [pem_pos=0, final_scores=0.68676335,
final_local_scores=0.7501721, pem=-0.20299491,
similarity=27.580261]
6: 15820894 Dynamo_Sports_Club
    [pem_pos=6, final_scores=0.29710156,
final_local_scores=0.32391265, pem=-2.1610804,
similarity=25.508204]
7: 5551651 KK_Partizan
    [pem_pos=1, final_scores=0.39474672,
final_local_scores=0.37364656, pem=-1.689159,
similarity=24.42598]
8: 5604103 KK_Crvena_zvezda
    [pem_pos=5, final_scores=0.36927104,
final_local_scores=0.36189884, pem=-2.4627666,
similarity=28.282404]
9: 34244 Yugoslavia
    [pem_pos=0, final_scores=0.3476742,
final_local_scores=0.33292428, pem=-0.6536551,
similarity=17.579515]
10: 11867 Germany
    [pem_pos=0, final_scores=0.67348367,
final_local_scores=0.73036516, pem=-0.15529114,
similarity=26.729763]
11: 3450530 Treviso_Basket
    [pem_pos=4, final_scores=0.54380983,
final_local_scores=0.5887608, pem=-2.1818364,
similarity=33.524673]
12: 14532 Italy
    [pem_pos=0, final_scores=0.562429,
final_local_scores=0.60858995, pem=-0.20612453,
similarity=23.372189]
13: 5551651 KK_Partizan
    [pem_pos=1, final_scores=0.3943116,
```

```
final_local_scores=0.37317127, pem=-1.689159,
similarity=24.411798]
14: 3450530 Treviso_Basket
    [pem_pos=4, final_scores=0.49928504,
final_local_scores=0.54012424, pem=-2.1818364,
similarity=32.07326]
15: 5551651 KK_Partizan
    [pem_pos=1, final_scores=0.29736346,
final_local_scores=0.26727018, pem=-1.689159,
similarity=21.251497]

Sample: b'814_SQUASH'
SQUASH - HONG KONG OPEN SECOND ROUND RESULTS . [1 1 HONG KONG
1] 1996-08-29 Second round results in the Hong Kong Open on
Thursday ( prefix number denotes seeding ) : 1 - [2 2 Jansher
Khan 2] ( [3 3 Pakistan 3] ) beat Simon Frenz ( [4 4 Germany 4
] ) 15-12 15-7 12-15 15-10 [5 5 Mark Cairns 5] ( [6 6 England
6] ) beat [7 7 Joseph Kneipp 7] ( [8 8 Australia 8] ) 8-15
15-12 15-8 15-8 Anthony Hill ( [9 9 Australia 9] ) beat Mir
Zaman Gul ( [10 10 Pakistan 10] ) 15-12 15-11 15-13 [11 11 Dan
Jenson 11] ( [12 12 Australia 12] ) beat 3 - [13 13 Brett Mar
tin 13] ( [14 14 Australia 14] ) 15-9 17-14 7-15 9-15 15-14 4
- [15 15 Peter Nicol 15] ( [16 16 Scotland 16] ) beat [17 17 J
onathon Power 17] ( [18 18 Canada 18] ) 15-10 15-9 15-9 7 - [1
9 19 Chris Walker 19] ( [20 20 England 20] ) beat [21 21 Amr S
habana 21] ( [22 22 Egypt 22] ) 15-13 15-10 15-6 [23 23 Derek
Ryan 23] ( [24 24 Ireland 24] ) beat [25 25 Paul Johnson 25]
( [26 26 England 26] ) 10-15 15-5 12-15 15-12 15-11 2 - [27 27
Rodney Eyles 27] ( [28 28 Australia 28] ) beat Zubair Jahan
Khan ( [29 29 Pakistan 29] ) 15-10 15-8 9-15 13-15 15-4 .
1: 13404 Hong_Kong
2: 2525758 Jansher_Khan
3: 23235 Pakistan
4: 11867 Germany
5: 7555328 Mark_Cairns_(squash_player)
6: 9316 England
7: 6849383 Joseph_Kneipp
8: 4689264 Australia
9: 4689264 Australia
10: 23235 Pakistan
11: 7447334 Dan_Jenson
12: 4689264 Australia
13: 3428114 Brett_Martin
14: 4689264 Australia
15: 2947276 Peter_Nicol
16: 26994 Scotland
17: 1529068 Jonathon_Power
18: 5042916 Canada
19: 3672106 Chris_Walker_(squash_player)
20: 9316 England
21: 3451400 Amr_Shabana
22: 8087628 Egypt
```

```
23: 24449519 Derek_Ryan_(squash_player)
24: 14560 Republic_of_Ireland
25: 7434625 Paul_Johnson_(squash_player)
26: 9316 England
27: 3428311 Rodney_Eyles
28: 4689264 Australia
29: 23235 Pakistan
1: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.51229423,
final_local_scores=0.51328564, pem=-0.22402541,
similarity=20.625465]
2: 2525758 Jansher_Khan
    [pem_pos=0, final_scores=0.53379184,
final_local_scores=0.645645, pem=0.0, similarity=23.357048]
3: 23235 Pakistan
    [pem_pos=0, final_scores=0.6430481,
final_local_scores=0.67320406, pem=-0.21944217,
similarity=25.372826]
4: 11867 Germany
    [pem_pos=0, final_scores=0.83254474,
final_local_scores=0.88765436, pem=-0.15529114,
similarity=31.423586]
5: 7555328 Mark_Cairns_(squash_player)
    [pem_pos=0, final_scores=0.5736233,
final_local_scores=0.68573767, pem=-0.18232158,
similarity=25.544987]
6: 9316 England
    [pem_pos=0, final_scores=0.5373752,
final_local_scores=0.5263044, pem=-0.19053856,
similarity=20.831865]
7: 6849383 Joseph_Kneipp
    [pem_pos=0, final_scores=0.70045054,
final_local_scores=0.81018335, pem=0.0, similarity=28.2672]
8: 4689264 Australia
    [pem_pos=0, final_scores=0.805955,
final_local_scores=0.808532, pem=-0.17168477,
similarity=29.15157]
9: 4689264 Australia
    [pem_pos=0, final_scores=0.73366815,
final_local_scores=0.72956955, pem=-0.17168477,
similarity=26.795174]
10: 23235 Pakistan
    [pem_pos=0, final_scores=0.6340806,
final_local_scores=0.66340846, pem=-0.21944217,
similarity=25.080505]
11: 7447334 Dan_Jenson
    [pem_pos=0, final_scores=0.65872794,
final_local_scores=0.8059559, pem=0.0, similarity=28.141045]
12: 4689264 Australia
    [pem_pos=0, final_scores=0.7815487,
final_local_scores=0.78187174, pem=-0.17168477,
similarity=28.355974]
```

```
13: 3428114 Brett_Martin
    [pem_pos=0, final_scores=0.734496,
final_local_scores=0.87065583, pem=0.0, similarity=30.07182]
14: 4689264 Australia
    [pem_pos=0, final_scores=0.7638696,
final_local_scores=0.76256007, pem=-0.17168477,
similarity=27.779676]
15: 2947276 Peter_Nicol
    [pem_pos=0, final_scores=0.6963464,
final_local_scores=0.7879952, pem=0.0, similarity=27.605062]
16: 26994 Scotland
    [pem_pos=0, final_scores=0.4499212,
final_local_scores=0.45130104, pem=-0.22905664,
similarity=18.803082]
17: 1529068 Jonathon_Power
    [pem_pos=0, final_scores=0.59548247,
final_local_scores=0.6807318, pem=0.0, similarity=24.404106]
18: 5042916 Canada
    [pem_pos=0, final_scores=0.76342756,
final_local_scores=0.7798701, pem=-0.11662793,
similarity=27.996832]
19: 3672106 Chris_Walker_(squash_player)
    [pem_pos=4, final_scores=0.6248537,
final_local_scores=0.7580133, pem=-1.4941665,
similarity=34.83585]
20: 9316 England
    [pem_pos=0, final_scores=0.51000416,
final_local_scores=0.49640572, pem=-0.19053856,
similarity=19.939629]
21: 3451400 Amr_Shabana
    [pem_pos=0, final_scores=0.58213526,
final_local_scores=0.7104953, pem=0.0, similarity=25.292309]
22: 8087628 Egypt
    [pem_pos=0, final_scores=0.561897,
final_local_scores=0.5967123, pem=-0.27480945,
similarity=23.391254]
23: 24449519 Derek_Ryan_(squash_player)
    [pem_pos=1, final_scores=0.4815898,
final_local_scores=0.60611445, pem=-1.3862944,
similarity=29.71626]
24: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.5338318,
final_local_scores=0.5635093, pem=-1.5040284,
similarity=29.085094]
25: 7434625 Paul_Johnson_(squash_player)
    [pem_pos=9, final_scores=0.4939618,
final_local_scores=0.6158067, pem=-2.962828,
similarity=38.578934]
26: 9316 England
    [pem_pos=0, final_scores=0.549738,
final_local_scores=0.5398089, pem=-0.19053856,
similarity=21.234867]
```

```
27: 3428311 Rodney_Eyles
    [pem_pos=0, final_scores=0.67090374,
final_local_scores=0.7792447, pem=0.0, similarity=27.34393]
28: 4689264 Australia
    [pem_pos=0, final_scores=0.75935787,
final_local_scores=0.75763166, pem=-0.17168477,
similarity=27.632603]
29: 23235 Pakistan
    [pem_pos=0, final_scores=0.6311771,
final_local_scores=0.66023684, pem=-0.21944217,
similarity=24.985857]
```

Sample: b'815_SOCCER'

SOCCER - RESULTS OF [1 1 SOUTH KOREAN 1] PRO-SOCCER GAMES . [2 2 SEOUL 2] 1996-08-29 Results of [3 3 South Korean 3] pro-soccer games played on Wednesday . Chonan 4 Anyang 1 (halftime 1-0) [4 4 Suwon 4] 4 Pusan 0 (halftime 2-0) Standings after games played on Wednesday (tabulate under - won , drawn , lost , goals for , goals against , points) : W D L G / F G / A P Chonan 3 0 1 13 10 9 Puchon 2 1 0 4 0 7 [5 5] Suwon 5] 1 3 0 7 3 6 [6 6 Pohang 6] 1 1 1 8 8 4 [7 7 Ulsan 7] 1 0 1 6 6 3 Anyang 0 3 1 6 9 3 Chonnam 0 2 1 4 5 2 Pusan 0 2 1 3 7 2 Chonbuk 0 0 2 2 5 0

1: 27019 South_Korea
2: 19159283 Seoul
3: 27019 South_Korea
4: 1054817 Suwon_Samsung_Bluewings
5: 1054817 Suwon_Samsung_Bluewings
6: 1056366 Pohang_Steelers
7: 1056425 Ulsan_Hyundai_FC
1: 27019 South_Korea
 [pem_pos=0, final_scores=0.46784025,
final_local_scores=0.45830083, pem=0.0, similarity=17.766323]
2: 19159283 Seoul
 [pem_pos=0, final_scores=0.4812376,
final_local_scores=0.4924152, pem=-0.0378014,
similarity=18.989935]
3: 27019 South_Korea
 [pem_pos=0, final_scores=0.5112371,
final_local_scores=0.50570536, pem=0.0, similarity=19.18097]
4: 1054817 Suwon_Samsung_Bluewings
 [pem_pos=2, final_scores=0.5246186,
final_local_scores=0.49959463, pem=-1.0623775,
similarity=24.775986]
5: 1054817 Suwon_Samsung_Bluewings
 [pem_pos=2, final_scores=0.52270555,
final_local_scores=0.4975049, pem=-1.0623775,
similarity=24.713625]
6: 1056366 Pohang_Steelers
 [pem_pos=2, final_scores=0.6174001,
final_local_scores=0.62669593, pem=-1.0587178,
similarity=28.549042]

```
7: 1056425 Ulsan_Hyundai_FC
    [pem_pos=3, final_scores=0.5551453,
final_local_scores=0.54839987, pem=-1.7336445,
similarity=29.882887]

Sample: b'816_BASEBALL'
BASEBALL - RESULTS OF [1 1 S. KOREAN 1] PRO-BASEBALL GAMES .
[2 2 SEOUL 2] 1996-08-29 Results of [3 3 South Korean 3] pro-
baseball games played on Wednesday . LG 5 0B 1 0B 4 LG 3
Ssangbangwool 12 [4 4 Hanwha 4] 0 [5 5 Hanwha 5] 12
Ssangbangwool 5 Lotte 4 Hyundai 0 Samsung 7 Haitai 1 Note -
LG and 0B , Ssangbangwool and [6 6 Hanwha 6] played two
games . Standings after games played on Wednesday ( tabulate
under won , drawn , lost , winning percentage , games behind
first place ) W D L PCT GB Haitai 63 2 42 .598 -
Ssangbangwool 59 2 48 .500 5 [7 7 Hanwha 7] 57 1 49 .537 6
1/2 Hyundai 56 5 48 .536 6 1/2 Samsung 48 5 55 .468 14 Lotte
45 6 53 .462 14 1/2 LG 45 5 59 .436 17 1/2 0B 42 6 61 .413 20
1: 27019 South_Korea
2: 19159283 Seoul
3: 27019 South_Korea
4: 4086558 Hanwha_Eagles
5: 4086558 Hanwha_Eagles
6: 4086558 Hanwha_Eagles
7: 4086558 Hanwha_Eagles
1: 27019 South_Korea
    [pem_pos=0, final_scores=0.5335104,
final_local_scores=0.5512212, pem=0.0, similarity=20.539253]
2: 19159283 Seoul
    [pem_pos=0, final_scores=0.46731955,
final_local_scores=0.49205774, pem=-0.0378014,
similarity=18.979267]
3: 27019 South_Korea
    [pem_pos=0, final_scores=0.49109608,
final_local_scores=0.5048899, pem=0.0, similarity=19.156635]
4: 4086558 Hanwha_Eagles
    [pem_pos=1, final_scores=0.71051335,
final_local_scores=0.69457924, pem=-0.6931472,
similarity=28.586784]
5: 4086558 Hanwha_Eagles
    [pem_pos=1, final_scores=0.71375984,
final_local_scores=0.69812554, pem=-0.6931472,
similarity=28.692614]
6: 4086558 Hanwha_Eagles
    [pem_pos=1, final_scores=0.52579314,
final_local_scores=0.49280047, pem=-0.6931472,
similarity=22.565304]
7: 4086558 Hanwha_Eagles
    [pem_pos=1, final_scores=0.6637714,
final_local_scores=0.6435208, pem=-0.6931472,
similarity=27.063099]
```

Sample: b'817_TENNIS'

TENNIS - THURSDAY 'S RESULTS FROM THE [1 1 U.S. OPEN 1] . [2 2 NEW YORK 2] 1996-08-29 Results of second round matches on Thursday in the [3 3 U.S. Open Tennis Championships 3] at the National Tennis Centre (prefix number denotes seeding) : Women 's singles [4 4 Anna Kournikova 4] ([5 5 Russia 5]) beat Natalia Baudone ([6 6 Italy 6]) 6-3 6-3 Rita Grande ([7 7 Italy 7]) beat [8 8 Tina Krizan 8] ([9 9 Slovenia 9]) 6-2 6-0 [10 10 Els Callens 10] ([11 11 Belgium 11]) beat Annabel Ellwood ([12 12 Australia 12]) 6-4 1-6 6-1 [13 13 Elena Likhovtseva 13] ([14 14 Russia 14]) beat Lila Osterloh ([15 15 U.S. 15]) 6-4 6-2 Sandra Dopfer ([16 16 Austria 16]) beat Nanne Dahlman ([17 17 Finland 17]) 2-6 6-2 6- 3 Men 's singles 13 - [18 18 Thomas Enqvist 18] ([19 19 Sweden 19]) beat [20 20 Guillaume Raoux 20] ([21 21 France 21]) 6-3 6- 2 6-3 [22 22 Sergi Bruguera 22] ([23 23 Spain 23]) beat [24 24 Michael Stich 24] ([25 25 Germany 25]) 6-3 6-2 6-4 [26 26 Jakob Hlasek 26] ([27 27 Switzerland 27]) beat [28 28 Alberto Berasategui 28] ([29 29 Spain 29]) 7-6 (7-5) 7-6 (9-7) 6-0 Women 's singles , second round 1 - [30 30 Steffi Graf 30] ([31 31 Germany 31]) beat [32 32 Karin Kschwendt 32] ([33 33 Austria 33]) 6-2 6-1 Naoko Kijimuta ([34 34 Japan 34]) beat [35 35 Alexandra Fusai 35] ([36 36 France 36]) 6-4 7-5 [37 37 Natasha Zvereva 37] ([38 38 Belarus 38]) beat [39 39 Ai Sugiyama 39] ([40 40 Japan 40]) 4-6 6-4 6-3 14 - [41 41 Barbara Paulus 41] ([42 42 Austria 42]) beat [43 43 Elena Wagner 43] ([44 44 Germany 44]) 7-5 7-6 (7-5) [45 45 Petra Langrova 45] ([46 46 Czech Republic 46]) beat [47 47 Naoko Sawamatsu 47] ([48 48 Japan 48]) 6- 4 3-6 7-5 17 - [49 49 Karina Habsudova 49] ([50 50 Slovakia 50]) beat [51 51 Nathalie Dechy 51] ([52 52 France 52]) 6-4 6-2 Women 's singles , second round 7 - [53 53 Jana Novotna 53] ([54 54 Czech Republic 54]) beat [55 55 Florencia Labat 55] ([56 56 Argentina 56]) 6-2 4-6 6-2 Men 's singles , second round 3 - [57 57 Thomas Muster 57] ([58 58 Austria 58]) beat Dirk Dier ([59 59 Germany 59]) 6-3 6-2 6-4 Pablo Campana ([60 60 Ecuador 60]) beat [61 61 Mark Knowles 61] ([62 62 Bahamas 62]) 7-6 (7-3) 3 - 6 6-3 6-7 (3-7) 6-3 [63 63 Jason Stoltenberg 63] ([64 64 Australia 64]) beat [65 65 Kenneth Carl sen 65] ([66 66 Denmark 66]) 6- 3 7-6 (7-1) 6-3 [67 67 Arnaud Boetsch 67] ([68 68 France 68]) beat [69 69 Magnus Gustafsson 69] ([70 70 Sweden 70]) 7-6 (8- 6) 6-3 6-1 Add Women 's singles , second round 16 - [71 71 Martina Hingis 71] ([72 72 Switzerland 72]) beat [73 73 Miriam Oremans 73] ([74 74 Netherlands 74]) 6-4 6-4 [75 75 Tami Whitlinger Jones 75] ([76 76 U.S. 76]) beat [77 77 Amy Frazier 77] ([78 78 U.S. 78]) 7-6 (7-3) 6-2 [79 79 Judith Wiesner 79] ([80 80 Austria 80]) beat [81 81 Debbie Graham 81] ([82 82 U.S. 82]) 6-2 7-5 Add Men 's singles , second round 6 - [83 83 Andre Agassi 83] ([84 84 U.S. 84]) beat [85 85 Leander Paes 85] ([86 86 India 86]) 3-6 6-4 6-1 6-0 [87 87 Javier Sanchez 87] ([88 88 Spain 88]) beat [89 89 Jim Grabb 89] ([90 90 U.S. 90]

) 6-2 7-6 (7-3) 2-6 6-3 [91 91 Hernan Gumy 91] ([92 92 Argentina 92]) beat [93 93 Jared Palmer 93] ([94 94 U.S. 94])
6-7 (5-7) 6-3 7-6 (7-4) 0-6 7-6 (7-1) : Women 's
singles , second round 3 - [95 95 Arantxa Sanchez Vicario 95]
([96 96 Spain 96]) beat [97 97 Nicole Arendt 97] ([98 98 U.S. 98]) 6-2 6-2 Men 's singles , second round [99 99 David Wheaton 99] ([100 100 U.S. 100]) beat Frederic Vitoux ([101 101 France 101]) 6-4 6-4 4-6 7-6 (7-4) [102 102 Jan Siemerink 102] ([103 103 Netherlands 103]) beat [104 104 Carlos Moyá 104] ([105 105 Spain 105]) 7-6 (7-2) 6-4 6-4
1: 82729 US_Open_(tennis)
2: 645042 New_York_City
3: 82729 US_Open_(tennis)
4: 890 Anna_Kournikova
5: 25391 Russia
6: 14532 Italy
7: 14532 Italy
8: 8858091 Tina_Križan
9: 27338 Slovenia
10: 1944281 Els_Callens
11: 3343 Belgium
12: 4689264 Australia
13: 2350072 Elena_Likhovtseva
14: 25391 Russia
15: 3434750 United_States
16: 26964606 Austria
17: 10577 Finland
18: 933709 Thomas_Enqvist
19: 5058739 Sweden
20: 3930324 Guillaume_Raoux
21: 5843419 France
22: 913625 Sergi_Bruguera
23: 26667 Spain
24: 979088 Michael_Stich
25: 11867 Germany
26: 1727962 Jakob_Hlasek
27: 26748 Switzerland
28: 1728009 Alberto_Berasategui
29: 26667 Spain
30: 164640 Steffi_Graf
31: 11867 Germany
32: 27388154 Karin_Kschwendt
33: 26964606 Austria
34: 15573 Japan
35: 4224688 Alexandra_Fusai
36: 5843419 France
37: 916019 Natasha_Zvereva
38: 3457 Belarus
39: 1514559 Ai_Sugiyama
40: 15573 Japan
41: 3128583 Barbara_Paulus
42: 26964606 Austria

43: 18330817 Elena_Pampoulova
44: 11867 Germany
45: 19609793 Petra_Langrová
46: 5321 Czech_Republic
47: 2631927 Naoko_Sawamatsu
48: 15573 Japan
49: 3062115 Karina_Habšudová
50: 26830 Slovakia
51: 1953990 Nathalie_Dechy
52: 5843419 France
53: 307050 Jana_Novotná
54: 5321 Czech_Republic
55: 3128595 Florencia_Labat
56: 18951905 Argentina
57: 166170 Thomas_Muster
58: 26964606 Austria
59: 11867 Germany
60: 9334 Ecuador
61: 2261491 Mark_Knowles
62: 3451 The_Bahamas
63: 2598440 Jason_Stoltenberg
64: 4689264 Australia
65: 1804967 Kenneth_Carlsen
66: 76972 Denmark
67: 3930278 Arnaud_Boetsch
68: 5843419 France
69: 2793059 Magnus_Gustafsson
70: 5058739 Sweden
71: 19047 Martina_Hingis
72: 26748 Switzerland
73: 2527510 Miriam_Oremans
74: 21148 Netherlands
75: 2783800 Tami_Whitlinger
76: 3434750 United_States
77: 2634301 Amy_Frazier
78: 3434750 United_States
79: 2805420 Judith_Wiesner
80: 26964606 Austria
81: 12556787 Debbie_Graham
82: 3434750 United_States
83: 595 Andre_Agassi
84: 3434750 United_States
85: 37491848 Leander_Paes
86: 14533 India
87: 420527 Javier_Sánchez
88: 26667 Spain
89: 1915197 Jim_Grabb
90: 3434750 United_States
91: 3929963 Hernán_Gumy
92: 18951905 Argentina
93: 3056059 Jared_Palmer
94: 3434750 United_States

```
95: 534624 Arantxa_Sánchez_Vicario
96: 26667 Spain
97: 4038024 Nicole_Arendt
98: 3434750 United_States
99: 1829027 David_Wheaton
100: 3434750 United_States
101: 5843419 France
102: 2876758 Jan_Siemerink
103: 21148 Netherlands
104: 669421 Carlos_Moyá
105: 26667 Spain
1: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.5145367,
final_local_scores=0.53233767, pem=-1.2370842,
similarity=26.703186]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.42528462,
final_local_scores=0.46012235, pem=-1.7906687,
similarity=27.558617]
3: 82729 US_Open_(tennis)
    [pem_pos=0, final_scores=0.57042474,
final_local_scores=0.5933868, pem=0.0, similarity=21.79756]
4: 890 Anna_Kournikova
    [pem_pos=0, final_scores=0.70086163,
final_local_scores=0.767617, pem=0.0, similarity=26.996935]
5: 25391 Russia
    [pem_pos=0, final_scores=0.87546045,
final_local_scores=0.9356882, pem=-0.20299491,
similarity=33.116432]
6: 14532 Italy
    [pem_pos=0, final_scores=0.74361306,
final_local_scores=0.7865008, pem=-0.20612453,
similarity=28.681404]
7: 14532 Italy
    [pem_pos=0, final_scores=0.7310677,
final_local_scores=0.772797, pem=-0.20612453,
similarity=28.272455]
8: 8858091 Tina_Križan
    [pem_pos=0, final_scores=0.84209985,
final_local_scores=0.97060484, pem=0.0, similarity=33.054497]
9: 27338 Slovenia
    [pem_pos=0, final_scores=0.5259388,
final_local_scores=0.5517647, pem=-0.27845925,
similarity=22.069777]
10: 1944281 Els_Callens
    [pem_pos=0, final_scores=0.56205714,
final_local_scores=0.6604697, pem=0.0, similarity=23.799446]
11: 3343 Belgium
    [pem_pos=0, final_scores=0.7235293,
final_local_scores=0.7233527, pem=-0.24180195,
similarity=26.990957]
12: 4689264 Australia
```

```
[pem_pos=0, final_scores=0.69217485,
final_local_scores=0.7328953, pem=-0.17168477,
similarity=26.89442]
13: 2350072 Elena_Likhovtseva
    [pem_pos=0, final_scores=0.7720602,
final_local_scores=0.8772126, pem=0.0, similarity=30.267487]
14: 25391 Russia
    [pem_pos=0, final_scores=0.84339577,
final_local_scores=0.90066236, pem=-0.20299491,
similarity=32.07119]
15: 3434750 United_States
    [pem_pos=0, final_scores=0.5221688,
final_local_scores=0.52849144, pem=-0.14075686,
similarity=20.626408]
16: 26964606 Austria
    [pem_pos=0, final_scores=0.65434974,
final_local_scores=0.6549851, pem=-0.25262418,
similarity=25.009584]
17: 10577 Finland
    [pem_pos=0, final_scores=0.6927393,
final_local_scores=0.7214203, pem=-0.2445894,
similarity=26.948448]
18: 933709 Thomas_Enqvist
    [pem_pos=0, final_scores=0.8276829,
final_local_scores=0.91359943, pem=0.0, similarity=31.353344]
19: 5058739 Sweden
    [pem_pos=0, final_scores=0.78262883,
final_local_scores=0.80878764, pem=-0.2135508,
similarity=29.386871]
20: 3930324 Guillaume_Raoux
    [pem_pos=0, final_scores=0.73249936,
final_local_scores=0.8302302, pem=0.0, similarity=28.865437]
21: 5843419 France
    [pem_pos=0, final_scores=0.7388143,
final_local_scores=0.74776584, pem=-0.16240834,
similarity=27.287739]
22: 913625 Sergi_Bruguera
    [pem_pos=0, final_scores=0.79337245,
final_local_scores=0.87274164, pem=0.0, similarity=30.134066]
23: 26667 Spain
    [pem_pos=0, final_scores=0.66954297,
final_local_scores=0.68443894, pem=-0.18857051,
similarity=25.540213]
24: 979088 Michael_Stich
    [pem_pos=0, final_scores=0.76098996,
final_local_scores=0.85129625, pem=0.0, similarity=29.494091]
25: 11867 Germany
    [pem_pos=0, final_scores=0.8454632,
final_local_scores=0.86769444, pem=-0.15529114,
similarity=30.827944]
26: 1727962 Jakob_Hlasek
    [pem_pos=0, final_scores=0.78371936,
```

```
final_local_scores=0.8910908, pem=0.0, similarity=30.68164]
27: 26748 Switzerland
    [pem_pos=0, final_scores=0.7085908,
final_local_scores=0.7151589, pem=-0.18598008,
similarity=26.44287]
28: 1728009 Alberto_Berasategui
    [pem_pos=0, final_scores=0.6545129,
final_local_scores=0.72723573, pem=0.0, similarity=25.791878]
29: 26667 Spain
    [pem_pos=0, final_scores=0.6350859,
final_local_scores=0.6467998, pem=-0.18857051,
similarity=24.416985]
30: 164640 Steffi_Graf
    [pem_pos=0, final_scores=0.8983926,
final_local_scores=0.95894367, pem=0.0, similarity=32.706505]
31: 11867 Germany
    [pem_pos=0, final_scores=0.88952726,
final_local_scores=0.9158277, pem=-0.15529114,
similarity=32.264336]
32: 27388154 Karin_Kschwendt
    [pem_pos=0, final_scores=0.8615089,
final_local_scores=0.9802919, pem=0.0, similarity=33.34358]
33: 26964606 Austria
    [pem_pos=0, final_scores=0.65309757,
final_local_scores=0.6536172, pem=-0.25262418,
similarity=24.968763]
34: 15573 Japan
    [pem_pos=0, final_scores=0.79206735,
final_local_scores=0.8489161, pem=-0.13549899,
similarity=30.159925]
35: 4224688 Alexandra_Fusai
    [pem_pos=0, final_scores=0.86705226,
final_local_scores=0.98639816, pem=0.0, similarity=33.525803]
36: 5843419 France
    [pem_pos=0, final_scores=0.7016419,
final_local_scores=0.70716065, pem=-0.16240834,
similarity=26.075998]
37: 916019 Natasha_Zvereva
    [pem_pos=0, final_scores=0.7415522,
final_local_scores=0.83622974, pem=0.0, similarity=29.044476]
38: 3457 Belarus
    [pem_pos=0, final_scores=0.58867663,
final_local_scores=0.6066901, pem=-0.4220667,
similarity=24.48982]
39: 1514559 Ai_Sugiyama
    [pem_pos=0, final_scores=0.74833965,
final_local_scores=0.8373439, pem=0.0, similarity=29.077724]
40: 15573 Japan
    [pem_pos=0, final_scores=0.72047645,
final_local_scores=0.77071387, pem=-0.13549899,
similarity=27.826216]
41: 3128583 Barbara_Paulus
```

[pem_pos=0, final_scores=0.5924818,
final_local_scores=0.7093244, pem=0.0, similarity=25.257368]
42: 26964606 Austria
[pem_pos=0, final_scores=0.6490113,
final_local_scores=0.6491536, pem=-0.25262418,
similarity=24.83556]
43: 18330817 Elena_Pampoulova
[pem_pos=0, final_scores=0.66064936,
final_local_scores=0.73912805, pem=0.0, similarity=26.146769]
44: 11867 Germany
[pem_pos=0, final_scores=0.85638005,
final_local_scores=0.8796194, pem=-0.15529114,
similarity=31.18381]
45: 19609793 Petra_Langrová
[pem_pos=0, final_scores=0.60016227,
final_local_scores=0.716907, pem=0.0, similarity=25.483648]
46: 5321 Czech_Republic
[pem_pos=0, final_scores=0.6254341,
final_local_scores=0.6287774, pem=-0.27794555,
similarity=24.365196]
47: 2631927 Naoko_Sawamatsu
[pem_pos=0, final_scores=0.5745133,
final_local_scores=0.68055123, pem=0.0, similarity=24.39872]
48: 15573 Japan
[pem_pos=0, final_scores=0.71249944,
final_local_scores=0.7620002, pem=-0.13549899,
similarity=27.566183]
49: 3062115 Karina_Habšudová
[pem_pos=0, final_scores=0.71169907,
final_local_scores=0.8638964, pem=0.0, similarity=29.870106]
50: 26830 Slovakia
[pem_pos=0, final_scores=0.66357213,
final_local_scores=0.66995627, pem=-0.36602598,
similarity=26.073051]
51: 1953990 Nathalie_Dechy
[pem_pos=0, final_scores=0.7178531,
final_local_scores=0.8162829, pem=0.0, similarity=28.449224]
52: 5843419 France
[pem_pos=0, final_scores=0.70421755,
final_local_scores=0.70997417, pem=-0.16240834,
similarity=26.15996]
53: 307050 Jana_Novotná
[pem_pos=0, final_scores=0.76555175,
final_local_scores=0.8497485, pem=0.0, similarity=29.447903]
54: 5321 Czech_Republic
[pem_pos=0, final_scores=0.6085023,
final_local_scores=0.61028206, pem=-0.27794555,
similarity=23.81326]
55: 3128595 Florencia_Labat
[pem_pos=0, final_scores=0.49860996,
final_local_scores=0.6094584, pem=0.0, similarity=22.277166]
56: 18951905 Argentina

```
[pem_pos=0, final_scores=0.6378917,
final_local_scores=0.6532954, pem=-0.30003807,
similarity=25.217005]
57: 166170 Thomas_Muster
    [pem_pos=0, final_scores=0.8754054,
final_local_scores=0.9427112, pem=0.0, similarity=32.222095]
58: 26964606 Austria
    [pem_pos=0, final_scores=0.610838,
final_local_scores=0.6074551, pem=-0.25262418,
similarity=23.591194]
59: 11867 Germany
    [pem_pos=0, final_scores=0.82223994,
final_local_scores=0.8423266, pem=-0.15529114,
similarity=30.070915]
60: 9334 Ecuador
    [pem_pos=0, final_scores=0.44233412,
final_local_scores=0.50309765, pem=-0.21404564,
similarity=20.267164]
61: 2261491 Mark_Knowles
    [pem_pos=0, final_scores=0.6094667,
final_local_scores=0.700679, pem=0.0, similarity=24.999372]
62: 3451 The_Bahamas
    [pem_pos=0, final_scores=0.4413767,
final_local_scores=0.48189133, pem=-0.35113433,
similarity=20.379835]
63: 2598440 Jason_Stoltenberg
    [pem_pos=0, final_scores=0.70634,
final_local_scores=0.79280204, pem=0.0, similarity=27.748507]
64: 4689264 Australia
    [pem_pos=0, final_scores=0.70263964,
final_local_scores=0.74432653, pem=-0.17168477,
similarity=27.23555]
65: 1804967 Kenneth_Carlsen
    [pem_pos=0, final_scores=0.71276563,
final_local_scores=0.80865645, pem=0.0, similarity=28.221634]
66: 76972 Denmark
    [pem_pos=0, final_scores=0.5979909,
final_local_scores=0.61974144, pem=-0.26762357,
similarity=24.039412]
67: 3930278 Arnaud_Boetsch
    [pem_pos=0, final_scores=0.66022533,
final_local_scores=0.7644005, pem=0.0, similarity=26.900948]
68: 5843419 France
    [pem_pos=0, final_scores=0.7434005,
final_local_scores=0.75277555, pem=-0.16240834,
similarity=27.437239]
69: 2793059 Magnus_Gustafsson
    [pem_pos=0, final_scores=0.7579314,
final_local_scores=0.8405122, pem=0.0, similarity=29.172274]
70: 5058739 Sweden
    [pem_pos=0, final_scores=0.7603231,
final_local_scores=0.784422, pem=-0.2135508,
```

```
similarity=28.659752]
71: 19047 Martina_Hingis
    [pem_pos=0, final_scores=0.78363687,
final_local_scores=0.84640026, pem=0.0, similarity=29.347984]
72: 26748 Switzerland
    [pem_pos=0, final_scores=0.69542426,
final_local_scores=0.70077646, pem=-0.18598008,
similarity=26.013668]
73: 2527510 Miriam_Oremans
    [pem_pos=0, final_scores=0.7462769,
final_local_scores=0.8847286, pem=0.0, similarity=30.49178]
74: 21148 Netherlands
    [pem_pos=0, final_scores=0.6294388,
final_local_scores=0.637233, pem=-0.24940795,
similarity=24.462337]
75: 2783800 Tami_Whitlinger
    [pem_pos=0, final_scores=0.55164146,
final_local_scores=0.6601932, pem=0.0, similarity=23.791195]
76: 3434750 United_States
    [pem_pos=0, final_scores=0.4928329,
final_local_scores=0.49644643, pem=-0.14075686,
similarity=19.670122]
77: 2634301 Amy_Frazier
    [pem_pos=0, final_scores=0.6089365,
final_local_scores=0.6896325, pem=0.0, similarity=24.669722]
78: 3434750 United_States
    [pem_pos=0, final_scores=0.5156805,
final_local_scores=0.52140397, pem=-0.14075686,
similarity=20.414906]
79: 2805420 Judith_Wiesner
    [pem_pos=0, final_scores=0.6176282,
final_local_scores=0.73207146, pem=0.0, similarity=25.936186]
80: 26964606 Austria
    [pem_pos=0, final_scores=0.6461416,
final_local_scores=0.6460189, pem=-0.25262418,
similarity=24.742016]
81: 12556787 Debbie_Graham
    [pem_pos=0, final_scores=0.56731874,
final_local_scores=0.7013491, pem=0.0, similarity=25.01937]
82: 3434750 United_States
    [pem_pos=0, final_scores=0.5065426,
final_local_scores=0.5114222, pem=-0.14075686,
similarity=20.117031]
83: 595 Andre_Agassi
    [pem_pos=0, final_scores=0.82587546,
final_local_scores=0.8718254, pem=0.0, similarity=30.106722]
84: 3434750 United_States
    [pem_pos=0, final_scores=0.5078437,
final_local_scores=0.51284343, pem=-0.14075686,
similarity=20.159443]
85: 37491848 Leander_Paes
    [pem_pos=0, final_scores=0.63702834,
```

```
final_local_scores=0.7126531, pem=0.0, similarity=25.356703]
86: 14533 India
    [pem_pos=0, final_scores=0.59170383,
final_local_scores=0.6638479, pem=-0.13622136,
similarity=24.641052]
87: 420527 Javier_Sánchez
    [pem_pos=0, final_scores=0.58917385,
final_local_scores=0.67962605, pem=0.0, similarity=24.371109]
88: 26667 Spain
    [pem_pos=0, final_scores=0.6202484,
final_local_scores=0.630592, pem=-0.18857051,
similarity=23.933311]
89: 1915197 Jim_Grabb
    [pem_pos=0, final_scores=0.72234017,
final_local_scores=0.8275116, pem=0.0, similarity=28.78431]
90: 3434750 United_States
    [pem_pos=0, final_scores=0.5142859,
final_local_scores=0.5198806, pem=-0.14075686,
similarity=20.369446]
91: 3929963 Hernán_Gumy
    [pem_pos=0, final_scores=0.71032685,
final_local_scores=0.817528, pem=0.0, similarity=28.48638]
92: 18951905 Argentina
    [pem_pos=0, final_scores=0.69399464,
final_local_scores=0.7145793, pem=-0.30003807,
similarity=27.045837]
93: 3056059 Jared_Palmer
    [pem_pos=0, final_scores=0.7333531,
final_local_scores=0.8605435, pem=0.0, similarity=29.770046]
94: 3434750 United_States
    [pem_pos=0, final_scores=0.5138264,
final_local_scores=0.5193786, pem=-0.14075686,
similarity=20.354464]
95: 534624 Arantxa_Sánchez_Vicario
    [pem_pos=0, final_scores=0.6525852,
final_local_scores=0.71580386, pem=0.0, similarity=25.450727]
96: 26667 Spain
    [pem_pos=0, final_scores=0.6494928,
final_local_scores=0.66253716, pem=-0.18857051,
similarity=24.88662]
97: 4038024 Nicole_Arendt
    [pem_pos=0, final_scores=0.81382686,
final_local_scores=0.9479751, pem=0.0, similarity=32.37918]
98: 3434750 United_States
    [pem_pos=0, final_scores=0.5107658,
final_local_scores=0.5160354, pem=-0.14075686,
similarity=20.254694]
99: 1829027 David_Wheaton
    [pem_pos=0, final_scores=0.5788112,
final_local_scores=0.65488327, pem=0.0, similarity=23.632736]
100: 3434750 United_States
    [pem_pos=0, final_scores=0.49657887,
```

```

final_local_scores=0.5005383, pem=-0.14075686,
similarity=19.79223]
101: 5843419 France
    [pem_pos=0, final_scores=0.7330689,
final_local_scores=0.7414898, pem=-0.16240834,
similarity=27.100449]
102: 2876758 Jan_Siemerink
    [pem_pos=0, final_scores=0.65691775,
final_local_scores=0.75304735, pem=0.0, similarity=26.562147]
103: 21148 Netherlands
    [pem_pos=0, final_scores=0.5787724,
final_local_scores=0.5818877, pem=-0.24940795,
similarity=22.810722]
104: 669421 Carlos_Moyá
    [pem_pos=0, final_scores=0.68295074,
final_local_scores=0.7452748, pem=0.0, similarity=26.330198]
105: 26667 Spain
    [pem_pos=0, final_scores=0.6127998,
final_local_scores=0.6224556, pem=-0.18857051,
similarity=23.690506]

```

Sample: b'818_BASEBALL'

BASEBALL - [1 1 MAJOR LEAGUE 1] STANDINGS AFTER WEDNESDAY 'S GAMES . [2 2 NEW YORK 2] 1996-08-29 [3 3 Major League Baseball 3] standings after games played on Wednesday (tabulate under won , lost , winning percentage and games behind) : [4 4 AMERICAN LEAGUE EASTERN DIVISION 4] W L PCT GB [1 1 NEW YORK 1] 74 58 .561 - [5 5 BALTIMORE 5] 70 62 .530 4 [6 6 BOSTON 6] 69 65 .515 6 [7 7 TORONTO 7] 63 71 .470 12 [8 8 DETROIT 8] 47 86 .353 27 1/2 [9 9 CENTRAL DIVISION 9] [10 10 CLEVELAND 10] 80 53 .602 - 2 CHICAGO 71 64 .526 10 [11 11 MINNESOTA 11] 66 67 .496 14 [12 12 MILWAUKEE 12] 64 70 .478 16 1/2 [13 13 KANSAS CITY 13] 61 73 .455 19 1/2 [14 14 WESTERN DIVISION 14] [15 15 TEXAS 15] 75 58 .564 - [16 16 SEATTLE 16] 69 63 .523 5 1/2 [17 17 OAKLAND 17] 64 72 .471 12 1/2 3 CALIFORNIA 61 72 .459 14 THURSDAY , AUGUST 29TH SCHEDULE [18 18 KANSAS CITY 18] AT [19 19 DETROIT 19] [20 20 MINNESOTA 20] AT [21 21 MILWAUKEE 21] [22 22 NEW YORK 22] AT 4 CALIFORNIA [23 23 BALTIMORE 23] AT [24 24 SEATTLE 24] [25 25 NATIONAL LEAGUE EASTERN DIVISION 25] W L PCT GB [26 26 ATLANTA 26] 82 49 .626 - [27 27 MONTREAL 27] 71 60 .542 11 5 FLORIDA 63 70 .474 20 [2 6 NEW YORK 2] 59 74 .444 24 [28 28 PHILADELPHIA 28] 54 80 .403 29 1/2 [29 29 CENTRAL DIVISION 29] [30 30 HOUSTON 30] 72 62 .537 - ST LOUIS 69 64 .519 2 1/2 [31 31 CINCINNATI 31] 65 67 .492 6 7 CHICAGO 64 66 .492 6 [32 32 PITTSBURGH 32] 56 76 .424 15 [33 33 WESTERN DIVISION 33] [34 34 SAN DIEGO 34] 74 60 .552 - 8 LOS ANGELES 71 61 .538 2 [35 35 COLORADO 35] 70 64 .522 4 [36 36 SAN FRANCISCO 36] 57 74 .435 15 1/2 THURSDAY , AUGUST 29TH SCHEDULE [37 37 SAN DIEGO 37] AT [3 9 NEW YORK 3] 10 CHICAGO AT [38 38 HOUSTON 38] [39 39 CINCINNATI 39] AT [40 40 COLORADO 40] [41 41 ATLANTA 41] AT [42 42 PITTSBURGH 42] 11 LOS ANGELES AT [43 43 MONTREAL 43] 12 FLORIDA AT ST LOUIS

```
1: 38776 Major_League_Baseball
2: 645042 New_York_City
3: 38776 Major_League_Baseball
4: 2004582 American_League_East
5: 4340 Baltimore_Orioles
6: 4339 Boston_Red_Sox
7: 30847 Toronto_Blue_Jays
8: 8579 Detroit_Tigers
9: 2456377 National_League_Central
10: 6652 Cleveland_Indians
11: 20050 Minnesota_Twins
12: 20152 Milwaukee_Brewers
13: 16883 Kansas_City_Royals
14: 2456474 National_League_West
15: 30857 Texas_Rangers_(baseball)
16: 27660 Seattle_Mariners
17: 22523 Oakland_Athletics
18: 16883 Kansas_City_Royals
19: 8579 Detroit_Tigers
20: 20050 Minnesota_Twins
21: 20152 Milwaukee_Brewers
22: 4848143 New_York_Yankees
23: 4340 Baltimore_Orioles
24: 27660 Seattle_Mariners
25: 2456294 National_League_East
26: 2140 Atlanta_Braves
27: 20153 Montreal_Expos
28: 23741 Philadelphia_Phillies
29: 2456377 National_League_Central
30: 13894 Houston_Astros
31: 6671 Cincinnati_Reds
32: 24270 Pittsburgh_Pirates
33: 2456474 National_League_West
34: 28417 San_Diego_Padres
35: 6669 Colorado_Rockies
36: 28416 San_Francisco_Giants
37: 28417 San_Diego_Padres
38: 13894 Houston_Astros
39: 6671 Cincinnati_Reds
40: 6669 Colorado_Rockies
41: 2140 Atlanta_Braves
42: 24270 Pittsburgh_Pirates
43: 20153 Montreal_Expos
1: 645042 New_York_City
    [winner: 4848143 New_York_Yankees, pem_pos=11,
final_scores=0.17992459, final_local_scores=0.08310217,
pem=-5.0692368, similarity=34.136944]
    gt: 645042 New_York_City, pem_pos=1,
final_scores=0.11473657, final_local_scores=0.15194172,
pem=-1.7906687, similarity=18.36189 ]
2: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
```

```
final_scores=-0.7210265, final_local_scores=-0.81031376,
pem=-4.591283, similarity=4.8764434|
    RECALL MISS: 5945 Chicago_White_Sox]
3: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 7590031 California_Golden_Bears_football,
pem_pos=1, final_scores=-0.90189976,
final_local_scores=-0.98998415, pem=-4.4771376,
similarity=-1.1060177|
    RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
4: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 5407 California, pem_pos=0,
final_scores=-0.6115518, final_local_scores=-0.6272217,
pem=-0.08662026, similarity=-14.15678|
    RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
5: 11278 Miami_Marlins
    [winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.74642867, final_local_scores=-0.8172681,
pem=-4.620281, similarity=4.8266115|
    RECALL MISS: 11278 Miami_Marlins]
6: 645042 New_York_City
    [winner: 4848143 New_York_Yankees, pem_pos=11,
final_scores=0.32186586, final_local_scores=0.23815137,
pem=-5.0692368, similarity=38.763924|
    gt: 645042 New_York_City, pem_pos=1,
final_scores=0.003283158, final_local_scores=0.030195802,
pem=-1.7906687, similarity=14.72875 ]
7: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.77743703, final_local_scores=-0.87193376,
pem=-4.591283, similarity=3.037579|
    RECALL MISS: 5945 Chicago_White_Sox]
8: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 18213 Los_Angeles_Dodgers, pem_pos=15,
final_scores=-0.11726761, final_local_scores=-0.2431277,
pem=-4.7524934, similarity=22.679096|
    RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
9: 645042 New_York_City
    [winner: 4848143 New_York_Yankees, pem_pos=11,
final_scores=0.25980473, final_local_scores=0.17035908,
pem=-5.0692368, similarity=36.740864|
    gt: 645042 New_York_City, pem_pos=1,
final_scores=0.01886414, final_local_scores=0.04721564,
pem=-1.7906687, similarity=15.236656 ]
10: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.65133184, final_local_scores=-0.73418295,
pem=-4.591283, similarity=7.148339|
    RECALL MISS: 5945 Chicago_White_Sox]
11: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 18213 Los_Angeles_Dodgers, pem_pos=15,
final_scores=-0.104212776, final_local_scores=-0.22886729,
pem=-4.7524934, similarity=23.104654|
```

RECALL MISS: 1360083 Los_Angeles_Angels_of_Anheim]
12: 11278 Miami_Marlins
[winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.6961212, final_local_scores=-0.76231486,
pem=-4.620281, similarity=6.466525]
RECALL MISS: 11278 Miami_Marlins]
1: 38776 Major_League_Baseball
[pem_pos=0, final_scores=0.5868215,
final_local_scores=0.5391101, pem=-0.47792417,
similarity=22.77686]
2: 645042 New_York_City
[pem_pos=1, final_scores=0.36884373,
final_local_scores=0.42936963, pem=-1.7906687,
similarity=26.640894]
3: 38776 Major_League_Baseball
[pem_pos=0, final_scores=0.5373634,
final_local_scores=0.48508465, pem=0.0, similarity=18.565607]
4: 2004582 American_League_East
[pem_pos=0, final_scores=0.47329658,
final_local_scores=0.47148544, pem=0.0, similarity=18.159779]
5: 4340 Baltimore_Orioles
[pem_pos=3, final_scores=1.236383,
final_local_scores=1.2353702, pem=-4.046561,
similarity=62.961445]
6: 4339 Boston_Red_Sox
[pem_pos=5, final_scores=1.2117615,
final_local_scores=1.2128084, pem=-4.418931,
similarity=64.31317]
7: 30847 Toronto_Blue_Jays
[pem_pos=6, final_scores=0.66411126,
final_local_scores=0.6202364, pem=-4.422637,
similarity=46.64979]
8: 8579 Detroit_Tigers
[pem_pos=4, final_scores=1.2661514,
final_local_scores=1.268429, pem=-4.1401463,
similarity=64.456924]
9: 2456377 National_League_Central
[pem_pos=0, final_scores=0.58215,
final_local_scores=0.6000702, pem=-0.98951703,
similarity=27.378153]
10: 6652 Cleveland_Indians
[pem_pos=12, final_scores=1.2520014,
final_local_scores=1.2520602, pem=-4.191953,
similarity=64.25018]
11: 20050 Minnesota_Twins
[pem_pos=9, final_scores=1.076163,
final_local_scores=1.0631187, pem=-4.2906566,
similarity=59.148556]
12: 20152 Milwaukee_Brewers
[pem_pos=4, final_scores=0.5246034,
final_local_scores=0.46244663, pem=-3.40205,
similarity=36.39092]

```
13: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.80756444,
final_local_scores=0.7720692, pem=-3.5915844,
similarity=46.661392]
14: 2456474 National_League_West
    [pem_pos=1, final_scores=0.5648642,
final_local_scores=0.56844527, pem=-1.2952092,
similarity=28.096802]
15: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.2741536,
final_local_scores=1.2811301, pem=-4.928425,
similarity=69.122734]
16: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.74775976,
final_local_scores=0.7117203, pem=-3.8567436,
similarity=46.302437]
17: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.3035488,
final_local_scores=1.3102107, pem=-3.9188845,
similarity=64.50052]
18: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.81265026,
final_local_scores=0.7776246, pem=-3.5915844,
similarity=46.827175]
19: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.0934472,
final_local_scores=1.079776, pem=-4.1401463,
similarity=58.82715]
20: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=1.1800954,
final_local_scores=1.1766491, pem=-4.2906566,
similarity=62.536522]
21: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.56811446,
final_local_scores=0.50997573, pem=-3.40205,
similarity=37.809284]
22: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.46436745,
final_local_scores=0.39372426, pem=-5.0692368,
similarity=43.40653]
23: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.0862049,
final_local_scores=1.0713234, pem=-4.046561,
similarity=58.065964]
24: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.68290824,
final_local_scores=0.6408798, pem=-3.8567436,
similarity=44.188416]
25: 2456294 National_League_East
    [pem_pos=0, final_scores=0.43150645,
final_local_scores=0.4240499, pem=0.0, similarity=16.744207]
26: 2140 Atlanta_Braves
```

```
[pem_pos=7, final_scores=1.1599355,
final_local_scores=1.1515188, pem=-4.136474,
similarity=60.948116]
27: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.79034716,
final_local_scores=0.75429857, pem=-4.5098605,
similarity=51.12481]
28: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.286205,
final_local_scores=1.291523, pem=-4.5189333,
similarity=67.20599]
29: 2456377 National_League_Central
    [pem_pos=0, final_scores=0.58596563,
final_local_scores=0.6042382, pem=-0.98951703,
similarity=27.502535]
30: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.3476921,
final_local_scores=1.3582888, pem=-4.650927,
similarity=69.91623]
31: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.1182252,
final_local_scores=1.1075153, pem=-3.771247,
similarity=57.648804]
32: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.2817174,
final_local_scores=1.2842422, pem=-4.03702,
similarity=64.368004]
33: 2456474 National_League_West
    [pem_pos=1, final_scores=0.58004016,
final_local_scores=0.5850226, pem=-1.2952092,
similarity=28.591503]
34: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.8655618,
final_local_scores=0.83651626, pem=-3.9307637,
similarity=50.42913]
35: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.1874437,
final_local_scores=1.1909082, pem=-4.482383,
similarity=64.004684]
36: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.9185819,
final_local_scores=0.8883607, pem=-4.409061,
similarity=54.577328]
37: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.87344295,
final_local_scores=0.8451253, pem=-3.9307637,
similarity=50.686043]
38: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.2863226,
final_local_scores=1.2912519, pem=-4.650927,
similarity=67.91571]
39: 6671 Cincinnati_Reds
```

```
[pem_pos=4, final_scores=1.2158673,
final_local_scores=1.2141745, pem=-3.771247,
similarity=60.83173]
40: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=0.91446036,
final_local_scores=0.89271516, pem=-4.482383,
similarity=55.106007]
41: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=1.1992546,
final_local_scores=1.194469, pem=-4.136474,
similarity=62.22984]
42: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.2779131,
final_local_scores=1.2800865, pem=-4.03702,
similarity=64.24399]
43: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.54502845,
final_local_scores=0.48632514, pem=-4.5098605,
similarity=43.12795]
1: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.17992459,
final_local_scores=0.08310217, pem=-5.0692368,
similarity=34.136944]
2: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.32186586,
final_local_scores=0.23815137, pem=-5.0692368,
similarity=38.763924]
3: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.25980473,
final_local_scores=0.17035908, pem=-5.0692368,
similarity=36.740864]

Sample: b'819_BASEBALL'
BASEBALL - [1 1 YANKEES 1] BRAWL AND CONTINUE TO SLIDE . [2 2
SEATTLE 2] 1996-08-29 [3 3 Jay Buhner 3] hit a three-run
homer and former [4 4 Yankee 4] [5 5 Terry Mulholland 5]
allowed one run over seven innings as the [6 6 Seattle Marine
rs 6] completed a sweep of [7 7 New York 7] with a 10-2
victory in a game marred by a bench-clearing brawl .
Including the last three games of last October 's Divisional
Playoff Series , the [8 8 Mariners 8] have beaten the [9 9 Ya
nkees 9] 12 of the past 15 meetings overall and 14 of the
last 16 in the [10 10 Kingdome 10] . Five players were
ejected after [11 11 Yankees 11] ' outfielde Paul O'Neill
and [12 12 Seattle 12] catcher [13 13 John Marzano 13] got
into a fight after O'Neill had been brushed back . In [14 14 B
altimore 14] , [15 15 Don Wengert 15] threw a nine-hitter for
his first shutout and [16 16 Jose Herrera 16] had a two-run
double in a three-run fifth inning as the [17 17 Oakland Athl
etics 17] blanked the [18 18 Baltimore Orioles 18] 3-0 . [19 1
9 Wengert 19] ( 7-9 ) , who failed to record a shutout in his
previous 86 starts in either the minors or majors , did not
```

walk a batter and struck out three for [20 20 Oakland 20] . In [21 21 Chicago 21] , [22 22 James Baldwin 22] scattered five hits over seven scoreless innings and [23 23 Ray Durham 23] and [24 24 Ozzie Guillen 24] had RBI hits in the second inning as the [25 25 Chicago White Sox 25] blanked the [26 26 Milwaukee Brewers 26] 2-0 . [27 27 Baldwin 27] (10-4) struck out four and did not walk a batter for 1 Chicago , which won for only the fourth time in 15 games . [28 28 Dave Nilsson 28] had three hits for the [29 29 Brewers 29] . In [30 30 Kansas City 30] , [31 31 Jose Offerman 31] 's single with two out in the 12th inning scored [32 32 Johnny Damon 32] with the winning run and lifted the [33 33 Kansas City Royals 33] to a 4-3 victory over the [34 34 Texas Rangers 34] . [35 35 Rick Huisman 35] (1-1) allowed one hit and a walk in the 12th to post his first major-league win . In [36 36 Toronto 36] , [37 37 Pat Hentgen 37] tossed a five-hitter for his fifth consecutive complete game and three players drove in two runs apiece as the [38 38 Toronto Blue Jays 38] defeated the [39 39 Minnesota Twins 39] 6-1 for their ninth win in 11 games . [40 40 Hentgen 40] (17-7) surrendered just three doubles and a pair of singles in tossing his major-league leading ninth complete game . He walked three and struck out three in winning for the 10th time in his last 11 decisions . In [41 41 Detroit 41] , [42 42 Orel Hershiser 42] recorded his fourth straight win and [43 43 Albert Belle 43] snapped a sixth-inning tie with a grand slam as the [44 44 Cleveland Indians 44] completed a season sweep of the [45 45 Detroit Tigers 45] with a 9-3 victory . [46 46 Hershiser 46] (14-7) , who allowed three runs , eight hits and one walk with five strikeouts over seven innings , improved to 4-0 in his last six starts , including a pair of wins over [47 47 Detroit 47] in the last 11 days . At [48 48 California 48] , In [49 49 Seattle 49] ,

1: 4848143 New_York_Yankees
2: 11388236 Seattle
3: 385914 Jay_Buhner
4: 4848143 New_York_Yankees
5: 1918847 Terry_Mulholland
6: 27660 Seattle_Mariners
7: 4848143 New_York_Yankees
8: 27660 Seattle_Mariners
9: 4848143 New_York_Yankees
10: 634056 Kingdome
11: 4848143 New_York_Yankees
12: 27660 Seattle_Mariners
13: 6694505 John_Marzano
14: 26997138 Baltimore
15: 6184750 Don_Wengert
16: 960136 José_Herrera_(1990s_outfielder)
17: 22523 Oakland_Athletics
18: 4340 Baltimore_Orioles
19: 6184750 Don_Wengert

```
20: 22523 Oakland_Athletics
21: 6886 Chicago
22: 1974033 James_Baldwin_(baseball)
23: 933727 Ray_Durham
24: 540550 Ozzie_Guillén
25: 5945 Chicago_White_Sox
26: 20152 Milwaukee_Brewers
27: 1974033 James_Baldwin_(baseball)
28: 2271683 Dave_Nilsson
29: 20152 Milwaukee_Brewers
30: 17454 Kansas_City,_Missouri
31: 2183067 José_Offerman
32: 608740 Johnny_Damon
33: 16883 Kansas_City_Royals
34: 30857 Texas_Rangers_(baseball)
35: 12174734 Rick_Huisman
36: 64646 Toronto
37: 383370 Pat_Hentgen
38: 30847 Toronto_Blue_Jays
39: 20050 Minnesota_Twins
40: 383370 Pat_Hentgen
41: 8687 Detroit
42: 493988 Orel_Hershiser
43: 1251621 Albert_Belle
44: 6652 Cleveland_Indians
45: 8579 Detroit_Tigers
46: 493988 Orel_Hershiser
47: 8579 Detroit_Tigers
48: 5407 California
49: 11388236 Seattle
1: 5945 Chicago_White_Sox
    [winner: 6611 Chicago_Bears, pem_pos=15,
final_scores=-0.49373674, final_local_scores=-0.5838099,
pem=-4.8795676, similarity=13.203508]
    RECALL MISS: 5945 Chicago_White_Sox]
1: 4848143 New_York_Yankees
    [pem_pos=0, final_scores=0.49622625,
final_local_scores=0.43165666, pem=-0.4294726,
similarity=19.306747]
2: 11388236 Seattle
    [pem_pos=0, final_scores=0.49306,
final_local_scores=0.5407737, pem=-0.1054736,
similarity=20.801062]
3: 385914 Jay_Buhner
    [pem_pos=0, final_scores=0.5428496,
final_local_scores=0.570323, pem=0.0, similarity=21.10929]
4: 4848143 New_York_Yankees
    [pem_pos=1, final_scores=0.4079345,
final_local_scores=0.33521125, pem=-2.2829223,
similarity=26.50797]
5: 1918847 Terry_Mulholland
    [pem_pos=0, final_scores=0.84432393,
```

```
final_local_scores=0.8792439, pem=0.0, similarity=30.328106]
6: 27660 Seattle_Mariners
    [pem_pos=0, final_scores=0.46891803,
final_local_scores=0.40641463, pem=-0.008242749,
similarity=16.262762]
7: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.57415897,
final_local_scores=0.517176, pem=-5.0692368,
similarity=47.090572]
8: 27660 Seattle_Mariners
    [pem_pos=0, final_scores=0.45494944,
final_local_scores=0.39115608, pem=-0.1261296,
similarity=16.448503]
9: 4848143 New_York_Yankees
    [pem_pos=0, final_scores=0.52955383,
final_local_scores=0.46806198, pem=-0.4294726,
similarity=20.393156]
10: 634056 Kingdome
    [pem_pos=0, final_scores=0.38659358,
final_local_scores=0.40181023, pem=0.0, similarity=16.080532]
11: 4848143 New_York_Yankees
    [pem_pos=0, final_scores=0.5191468,
final_local_scores=0.45669383, pem=-0.4294726,
similarity=20.053905]
12: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.62772924,
final_local_scores=0.57989174, pem=-3.8567436,
similarity=42.36841]
13: 6694505 John_Marzano
    [pem_pos=0, final_scores=0.48366612,
final_local_scores=0.56352586, pem=0.0, similarity=20.906448]
14: 26997138 Baltimore
    [pem_pos=0, final_scores=0.4054659,
final_local_scores=0.44163585, pem=-0.19949137,
similarity=18.353872]
15: 6184750 Don_Wengert
    [pem_pos=0, final_scores=0.48741436,
final_local_scores=0.5594532, pem=0.0, similarity=20.784912]
16: 960136 José_Herrera_(1990s_outfielder)
    [pem_pos=2, final_scores=0.21975349,
final_local_scores=0.2724079, pem=-1.0986123,
similarity=18.19333]
17: 22523 Oakland_Athletics
    [pem_pos=0, final_scores=0.5387315,
final_local_scores=0.4808488, pem=-0.06842491,
similarity=18.811304]
18: 4340 Baltimore_Orioles
    [pem_pos=0, final_scores=0.444184,
final_local_scores=0.3772303, pem=-0.21811225,
similarity=16.533146]
19: 6184750 Don_Wengert
    [pem_pos=0, final_scores=0.6215679,
```

```
final_local_scores=0.7059955, pem=0.0, similarity=25.158028]
20: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=0.927311,
final_local_scores=0.90531284, pem=-3.9188845,
similarity=52.417557]
21: 6886 Chicago
    [pem_pos=0, final_scores=0.5161094,
final_local_scores=0.5529559, pem=-0.20658566,
similarity=21.714466]
22: 1974033 James_Baldwin_(baseball)
    [pem_pos=1, final_scores=0.5564798,
final_local_scores=0.5765087, pem=-1.6250951,
similarity=30.131403]
23: 933727 Ray_Durham
    [pem_pos=0, final_scores=0.671549,
final_local_scores=0.6991241, pem=0.0, similarity=24.95297]
24: 540550 Ozzie_Guillén
    [pem_pos=0, final_scores=0.5586166,
final_local_scores=0.5654703, pem=0.0, similarity=20.964474]
25: 5945 Chicago_White_Sox
    [pem_pos=0, final_scores=0.42916626,
final_local_scores=0.3592578, pem=-0.07886935,
similarity=15.239586]
26: 20152 Milwaukee_Brewers
    [pem_pos=0, final_scores=0.41233367,
final_local_scores=0.34370214, pem=-0.12176443,
similarity=15.008644]
27: 1974033 James_Baldwin_(baseball)
    [pem_pos=1, final_scores=0.6155506,
final_local_scores=0.64103454, pem=-1.6250951,
similarity=32.056984]
28: 2271683 Dave_Nilsson
    [pem_pos=0, final_scores=0.5846296,
final_local_scores=0.6428316, pem=0.0, similarity=23.27309]
29: 20152 Milwaukee_Brewers
    [pem_pos=0, final_scores=0.55229646,
final_local_scores=0.4965902, pem=-0.24813148,
similarity=20.258331]
30: 17454 Kansas_City,_Missouri
    [pem_pos=0, final_scores=0.32625705,
final_local_scores=0.33204806, pem=-0.8046767,
similarity=18.374647]
31: 2183067 José_Offerman
    [pem_pos=0, final_scores=0.48505414,
final_local_scores=0.5147991, pem=0.0, similarity=19.452345]
32: 608740 Johnny_Damon
    [pem_pos=0, final_scores=0.7193556,
final_local_scores=0.73588043, pem=0.0, similarity=26.049852]
33: 16883 Kansas_City_Royals
    [pem_pos=0, final_scores=0.5284247,
final_local_scores=0.46870548, pem=-0.029036561,
similarity=18.234726]
```

```
34: 30857 Texas_Rangers_(baseball)
    [pem_pos=0, final_scores=0.55320066,
final_local_scores=0.49732953, pem=-0.15069729,
similarity=19.750534]
35: 12174734 Rick_Huisman
    [pem_pos=0, final_scores=0.5099831,
final_local_scores=0.5888609, pem=0.0, similarity=21.662498]
36: 64646 Toronto
    [pem_pos=0, final_scores=0.6612244,
final_local_scores=0.73448277, pem=-0.12291032,
similarity=26.67655]
37: 383370 Pat_Hentgen
    [pem_pos=0, final_scores=0.6356342,
final_local_scores=0.66534746, pem=0.0, similarity=23.945007]
38: 30847 Toronto_Blue_Jays
    [pem_pos=0, final_scores=0.45352787,
final_local_scores=0.392572, pem=-0.018220792,
similarity=15.903931]
39: 20050 Minnesota_Twins
    [pem_pos=0, final_scores=0.4731596,
final_local_scores=0.41211724, pem=-0.052076127,
similarity=16.67131]
40: 383370 Pat_Hentgen
    [pem_pos=0, final_scores=0.38432568,
final_local_scores=0.39083105, pem=0.0, similarity=15.752891]
41: 8687 Detroit
    [pem_pos=0, final_scores=0.52593124,
final_local_scores=0.56134516, pem=-0.21700948,
similarity=22.021503]
42: 493988 Orel_Hershiser
    [pem_pos=0, final_scores=0.66729337,
final_local_scores=0.68363833, pem=0.0, similarity=24.490845]
43: 1251621 Albert_Belle
    [pem_pos=0, final_scores=0.62920564,
final_local_scores=0.66254985, pem=0.0, similarity=23.861523]
44: 6652 Cleveland_Indians
    [pem_pos=0, final_scores=0.60309535,
final_local_scores=0.5481537, pem=-0.08813761,
similarity=20.92702]
45: 8579 Detroit_Tigers
    [pem_pos=0, final_scores=0.5151519,
final_local_scores=0.45470482, pem=-0.08904677,
similarity=18.143261]
46: 493988 Orel_Hershiser
    [pem_pos=0, final_scores=0.64479536,
final_local_scores=0.6590627, pem=0.0, similarity=23.757458]
47: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=0.7753946,
final_local_scores=0.7389804, pem=-4.1401463,
similarity=48.657124]
48: 5407 California
    [pem_pos=0, final_scores=0.5654062,
```

```
final_local_scores=0.6481151, pem=-0.08662026,  
similarity=23.901815]  
49: 11388236 Seattle  
    [pem_pos=0, final_scores=0.5011668,  
final_local_scores=0.5496292, pem=-0.1054736,  
similarity=21.065327]
```

Sample: b'820_BASEBALL'

BASEBALL - [1 1 MAJOR LEAGUE 1] RESULTS WEDNESDAY . [2 2 NEW YORK 2] 1996-08-29 Results of [3 3 Major League 3] Baseball games played on Wednesday (home team in CAPS) : [4 4 National League 4] [5 5 COLORADO 5] 10 [6 6 Cincinnati 6] 9 [7 7 MONTREAL 7] 6 [8 8 Los Angeles 8] 5 [9 9 Atlanta 9] 9 [10 10 PITTSBURGH 10] 4 [11 11 San Diego 11] 3 [12 12 NEW YORK 12] 2 (10 innings) [13 13 HOUSTON 13] 5 1 Chicago 4 (11 innings) 2 Florida 3 ST LOUIS 2 (10 innings) [14 14 SAN FRANCISCO 14] 7 [15 15 Philadelphia 15] 6 [16 16 American League 16] [17 17 Cleveland 17] 9 [18 18 DETROIT 18] 3 3 CHICAGO 2 [19 19 Milwaukee 19] 0 [20 20 Oakland 20] 3 [21 21 BALTIMORE 21] 0 [22 22 TORONTO 22] 6 [23 23 Minnesota 23] 1 [24 24 KANSAS CITY 24] 4 [25 25 Texas 25] 3 (12 innings) [26 26 Boston 26] 7 4 CALIFORNIA 4 [27 27 SEATTLE 27] 10 [28 28 New York 28] 2

1: 38776 Major_League_Baseball
2: 645042 New_York_City
3: 38776 Major_League_Baseball
4: 64290 National_League
5: 6669 Colorado_Rockies
6: 6671 Cincinnati_Reds
7: 20153 Montreal_Expos
8: 18213 Los_Angeles_Dodgers
9: 2140 Atlanta_Braves
10: 24270 Pittsburgh_Pirates
11: 28417 San_Diego_Padres
12: 4848143 New_York_Yankees
13: 13894 Houston_Astros
14: 28416 San_Francisco_Giants
15: 23741 Philadelphia_Phillies
16: 64777 American_League
17: 6652 Cleveland_Indians
18: 8579 Detroit_Tigers
19: 20152 Milwaukee_Brewers
20: 22523 Oakland_Athletics
21: 4340 Baltimore_Orioles
22: 30847 Toronto_Blue_Jays
23: 20050 Minnesota_Twins
24: 16883 Kansas_City_Royals
25: 30857 Texas_Rangers_(baseball)
26: 4339 Boston_Red_Sox
27: 27660 Seattle_Mariners
28: 4848143 New_York_Yankees
1: 5945 Chicago_White_Sox
 [winner: 73126 Chicago_Blackhawks, pem_pos=8,

```
final_scores=-0.713653, final_local_scores=-0.7996244,
pem=-4.591283, similarity=5.195435|
    RECALL MISS: 5945 Chicago_White_Sox]
2: 11278 Miami_Marlins
    [winner: 70551 Florida_Panthers, pem_pos=5,
final_scores=-0.74408203, final_local_scores=-0.8113722,
pem=-4.620281, similarity=5.002555|
    RECALL MISS: 11278 Miami_Marlins]
3: 5945 Chicago_White_Sox
    [winner: 73126 Chicago_Blackhawks, pem_pos=8,
final_scores=-0.7936746, final_local_scores=-0.8870358,
pem=-4.591283, similarity=2.586905|
    RECALL MISS: 5945 Chicago_White_Sox]
4: 1360083 Los_Angeles_Angels_ofAnaheim
    [winner: 5407 California, pem_pos=0,
final_scores=-0.59230024, final_local_scores=-0.59994835,
pem=-0.08662026, similarity=-13.342888|
    RECALL MISS: 1360083 Los_Angeles_Angels_ofAnaheim]
1: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.55472654,
final_local_scores=0.50140977, pem=-0.47792417,
similarity=21.651806]
2: 645042 New_York_City
    [pem_pos=1, final_scores=0.28343445,
final_local_scores=0.35026443, pem=-1.7906687,
similarity=24.280237]
3: 38776 Major_League_Baseball
    [pem_pos=0, final_scores=0.5559283,
final_local_scores=0.5027225, pem=-0.47792417,
similarity=21.69098]
4: 64290 National_League
    [pem_pos=0, final_scores=0.6160385,
final_local_scores=0.585982, pem=-0.07431644,
similarity=21.98073]
5: 6669 Colorado_Rockies
    [pem_pos=11, final_scores=1.1236103,
final_local_scores=1.1213725, pem=-4.482383,
similarity=61.929596]
6: 6671 Cincinnati_Reds
    [pem_pos=4, final_scores=1.0750117,
final_local_scores=1.0588012, pem=-3.771247,
similarity=56.19508]
7: 20153 Montreal_Expos
    [pem_pos=8, final_scores=0.70805794,
final_local_scores=0.6645832, pem=-4.5098605,
similarity=48.447525]
8: 18213 Los_Angeles_Dodgers
    [pem_pos=15, final_scores=0.56566215,
final_local_scores=0.50271314, pem=-4.7524934,
similarity=44.936474]
9: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=0.94333667,
```

```
final_local_scores=0.91449326, pem=-4.136474,
similarity=53.874805]
10: 24270 Pittsburgh_Pirates
    [pem_pos=7, final_scores=1.1415321,
final_local_scores=1.1297243, pem=-4.03702,
similarity=59.756878]
11: 28417 San_Diego_Padres
    [pem_pos=6, final_scores=0.61840135,
final_local_scores=0.56563306, pem=-3.9307637,
similarity=42.345436]
12: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.52727765,
final_local_scores=0.46280414, pem=-5.0692368,
similarity=45.46801]
13: 13894 Houston_Astros
    [pem_pos=11, final_scores=1.3952781,
final_local_scores=1.410356, pem=-4.650927,
similarity=71.47002]
14: 28416 San_Francisco_Giants
    [pem_pos=3, final_scores=0.876937,
final_local_scores=0.8420226, pem=-4.409061,
similarity=53.194504]
15: 23741 Philadelphia_Phillies
    [pem_pos=10, final_scores=1.0629258,
final_local_scores=1.0458584, pem=-4.5189333,
similarity=59.874874]
16: 64777 American_League
    [pem_pos=0, final_scores=0.5514206,
final_local_scores=0.5205478, pem=0.0, similarity=19.623898]
17: 6652 Cleveland_Indians
    [pem_pos=12, final_scores=1.0987345,
final_local_scores=1.0827389, pem=-4.191953,
similarity=59.197292]
18: 8579 Detroit_Tigers
    [pem_pos=4, final_scores=1.078613,
final_local_scores=1.0616753, pem=-4.1401463,
similarity=58.28698]
19: 20152 Milwaukee_Brewers
    [pem_pos=4, final_scores=0.4067031,
final_local_scores=0.332753, pem=-3.40205,
similarity=32.520603]
20: 22523 Oakland_Athletics
    [pem_pos=7, final_scores=1.2990427,
final_local_scores=1.3034196, pem=-3.9188845,
similarity=64.29786]
21: 4340 Baltimore_Orioles
    [pem_pos=3, final_scores=1.1795343,
final_local_scores=1.1716902, pem=-4.046561,
similarity=61.061108]
22: 30847 Toronto_Blue_Jays
    [pem_pos=6, final_scores=0.78120595,
final_local_scores=0.747483, pem=-4.422637,
```

```
similarity=50.447086]
23: 20050 Minnesota_Twins
    [pem_pos=9, final_scores=0.84100944,
final_local_scores=0.8042138, pem=-4.2906566,
similarity=51.422314]
24: 16883 Kansas_City_Royals
    [pem_pos=5, final_scores=0.8839027,
final_local_scores=0.8541949, pem=-3.5915844,
similarity=49.112186]
25: 30857 Texas_Rangers_(baseball)
    [pem_pos=13, final_scores=1.2295916,
final_local_scores=1.231152, pem=-4.928425,
similarity=67.63129]
26: 4339 Boston_Red_Sox
    [pem_pos=5, final_scores=1.1824877,
final_local_scores=1.1796148, pem=-4.418931,
similarity=63.322605]
27: 27660 Seattle_Mariners
    [pem_pos=4, final_scores=0.750573,
final_local_scores=0.71459764, pem=-3.8567436,
similarity=46.3883]
28: 4848143 New_York_Yankees
    [pem_pos=11, final_scores=0.53285235,
final_local_scores=0.4688937, pem=-5.0692368,
similarity=45.649734]
```

Sample: b'821_TENNIS'

TENNIS - [1 1 MARTINEZ 1] GETS AGGRESSIVE AT [2 2 U.S. OPEN 2] . Larry Fine [3 3 NEW YORK 3] 1996-08-28 [4 4 Conchita Martinez 4] decided that when in [5 5 New York 5] , do as the [6 6 New Yorkers 6] do -- and the [7 7 Spaniard 7] 's new-found aggressiveness seems to have put her in the right frame of mind for the [8 8 U.S. Open 8] . The fourth-seeded [9 9 Spaniard 9] , who is tackling the world class traffic of [10 10 New York City 10] as a warm-up by driving to the tennis centre for her matches , ran over [11 11 France 11] 's [12 12 Nathalie Tauziat 12] 6-1 6-3 on Wednesday to take her place in the third round . " I 've been trying my whole career to be aggressive , " said the 24-year-old [13 13 Martinez 13] after crushing [14 14 Tauziat 14] in just over an hour . " What I 'm trying to do is be aggressive all the time , maybe go up to the net a few times like I did tonight . That would really help . " [15 15 Martinez 15] , the 1994 [16 16 Wimbledon 16] champion , used to struggle at the Open , but has come to terms with the noise , crowds and chaos . " There is a lot of things that can happen , " [17 17 Martinez 17] said about her early difficulties adjusting to tennis on the cement at [18 18 Flushing Meadows 18] . " Like traffic . We stay in [19 19 Manhattan 19] and it 's a long way to come . The crowds , they speak louder or they move . That does n't happen in other [20 20 Grand Slams 20] . That 's where the real champion wins . You have to concentrate for these two weeks . " It took [21 21

Martinez 21] four Opens to get as far as the quarters , and another four to match that . Last year , [22 22 **Martinez** 22] , who finished 1995 ranked second in the world , reached the semifinals before bowing out to [23 23 **Monica Seles** 23] . Now she feels she is in the swing of things . " I have my own car now and that helps , " said [24 24 **Martinez** 24] . " Sometimes the transportation they (the tournament) provide gets a little messed up . This time it did n't happen . I do the driving and I love it . It gets my adrenalin going , those taxi drivers . " We change lanes all the time in [25 25 **Barcelona** 25] . I 'm used to it and I like to drive fast . "

```
1: 784152 Conchita_Martínez
2: 82729 US_Open_(tennis)
3: 645042 New_York_City
4: 784152 Conchita_Martínez
5: 645042 New_York_City
6: 645042 New_York_City
7: 26667 Spain
8: 82729 US_Open_(tennis)
9: 26667 Spain
10: 645042 New_York_City
11: 5843419 France
12: 1028366 Nathalie_Tauziat
13: 784152 Conchita_Martínez
14: 1028366 Nathalie_Tauziat
15: 784152 Conchita_Martínez
16: 91195 The_Championships,_Wimbledon
17: 784152 Conchita_Martínez
18: 89050 Flushing_Meadows–Corona_Park
19: 45470 Manhattan
20: 197638 Grand_Slam_(tennis)
21: 784152 Conchita_Martínez
22: 784152 Conchita_Martínez
23: 300468 Monica_Séles
24: 784152 Conchita_Martínez
25: 4443 Barcelona
1: 784152 Conchita_Martínez
    [pem_pos=0, final_scores=1.0576711,
final_local_scores=1.073127, pem=0.0, similarity=36.113964]
2: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.5814027,
final_local_scores=0.58229256, pem=-1.2370842,
similarity=28.19394]
3: 645042 New_York_City
    [pem_pos=1, final_scores=0.3682714,
final_local_scores=0.4467554, pem=-1.7906687,
similarity=27.15972]
4: 784152 Conchita_Martínez
    [pem_pos=0, final_scores=0.9425259,
final_local_scores=0.9439171, pem=0.0, similarity=32.258083]
5: 645042 New_York_City
    [pem_pos=1, final_scores=0.45430082,
```

```
final_local_scores=0.54072946, pem=-1.7906687,
similarity=29.964092]
6: 645042 New_York_City
    [pem_pos=0, final_scores=0.40298885,
final_local_scores=0.48467898, pem=0.0, similarity=18.553501]
7: 26667 Spain
    [pem_pos=1, final_scores=0.2974993,
final_local_scores=0.38022256, pem=-0.7552963,
similarity=19.543732]
8: 82729 US_Open_(tennis)
    [pem_pos=1, final_scores=0.4791276,
final_local_scores=0.47057247, pem=-1.2370842,
similarity=24.85999]
9: 26667 Spain
    [pem_pos=3, final_scores=0.38484848,
final_local_scores=0.47563827, pem=-0.7552963,
similarity=22.391127]
10: 645042 New_York_City
    [pem_pos=0, final_scores=0.5010626,
final_local_scores=0.59180963, pem=0.0, similarity=21.750492]
11: 5843419 France
    [pem_pos=0, final_scores=0.59376436,
final_local_scores=0.6932957, pem=-0.16240834,
similarity=25.662241]
12: 1028366 Nathalie_Tauziat
    [pem_pos=0, final_scores=0.7549872,
final_local_scores=0.8086403, pem=0.0, similarity=28.221153]
13: 784152 Conchita_Martínez
    [pem_pos=0, final_scores=1.4067231,
final_local_scores=1.4544134, pem=0.0, similarity=47.49231]
14: 1028366 Nathalie_Tauziat
    [pem_pos=0, final_scores=0.94459385,
final_local_scores=1.0157568, pem=0.0, similarity=34.40192]
15: 784152 Conchita_Martínez
    [pem_pos=0, final_scores=1.5937036,
final_local_scores=1.6586611, pem=0.0, similarity=53.587467]
16: 91195 The_Championships,_Wimbledon
    [pem_pos=0, final_scores=0.488769,
final_local_scores=0.48888117, pem=-0.8505555,
similarity=23.304356]
17: 784152 Conchita_Martínez
    [pem_pos=0, final_scores=1.4313498,
final_local_scores=1.4813142, pem=0.0, similarity=48.295082]
18: 89050 Flushing_Meadows–Corona_Park
    [pem_pos=0, final_scores=0.21107106,
final_local_scores=0.2781749, pem=0.0, similarity=12.391007]
19: 45470 Manhattan
    [pem_pos=0, final_scores=0.2180727,
final_local_scores=0.2985185, pem=-0.13335879,
similarity=13.723327]
20: 197638 Grand_Slam_(tennis)
    [pem_pos=0, final_scores=0.4957927,
```

```
final_local_scores=0.49465173, pem=0.0, similarity=18.851107]
21: 784152 Conchita_Martínez
    [pem_pos=0, final_scores=1.5377263,
final_local_scores=1.5975144, pem=0.0, similarity=51.762726]
22: 784152 Conchita_Martínez
    [pem_pos=0, final_scores=1.5031365,
final_local_scores=1.5597303, pem=0.0, similarity=50.635178]
23: 300468 Monica_Seles
    [pem_pos=0, final_scores=0.7402454,
final_local_scores=0.7483503, pem=0.0, similarity=26.42198]
24: 784152 Conchita_Martínez
    [pem_pos=0, final_scores=1.4248309,
final_local_scores=1.4741933, pem=0.0, similarity=48.08258]
25: 4443 Barcelona
    [pem_pos=0, final_scores=0.5038306,
final_local_scores=0.56766015, pem=-0.37077636,
similarity=23.046164]
```

Sample: b'822_BASEBALL'

BASEBALL - [1 1 BRAVES 1] SIGN [2 2 NEAGLE 2] . [3 3 ATLANTA
3] 1996-08-28 The defending world champion [4 4 Atlanta Braves
4] , with the best record and best pitching in baseball ,
added another weapon Wednesday , acquiring [5 5 Denny Neagle 5
] , the winningest left-hander in the [6 6 National League 6]
, from the [7 7 Pittsburgh Pirates 7] . The [8 8 Pirates 8] ,
who conceded earlier this week they would be forced to trim
salary from next season 's payroll , received [9 9 Ron Wright
9] , a first baseman at Double-A 1 Greenville ; Corey
Pointer , a pitcher at Class-A [10 10 Eugene 10] , and a
player to be named . It was another stunning mid-season
acquisition for the [11 11 Braves 11] , who already have an
11-game lead in the [12 12 National League Eastern Division 12]
] . In the last 15 days , [13 13 Atlanta 13] has traded for
third baseman [14 14 Terry Pendleton 14] , claimed outfielder
[15 15 Luis Polonia 15] on waivers and called up minor-league
phenom [16 16 Andruw Jones 16] , all in preparation for their
fifth post-season .

```
1: 2140 Atlanta_Braves
2: 4212649 Denny_Neagle
3: 3138 Atlanta
4: 2140 Atlanta_Braves
5: 4212649 Denny_Neagle
6: 64290 National_League
7: 24270 Pittsburgh_Pirates
8: 24270 Pittsburgh_Pirates
9: 12142087 Ron_Wright_(baseball)
10: 1655051 Eugene_Emeralds
11: 2140 Atlanta_Braves
12: 2456294 National_League_East
13: 2140 Atlanta_Braves
14: 486632 Terry_Pendleton
15: 5158901 Luis_Polonia
```

```
16: 875934 Andruw_Jones
1: 1050189 Greenville_Braves
    [winner: 1050189 Greenville_Braves, pem_pos=0,
final_scores=-0.2132711, final_local_scores=-0.25467634,
pem=0.0, similarity=-3.5103374]
    gt: 1050189 Greenville_Braves, pem_pos=0,
final_scores=-0.2132711, final_local_scores=-0.25467634,
pem=0.0, similarity=-3.5103374 ]
1: 2140 Atlanta_Braves
    [pem_pos=0, final_scores=0.5431085,
final_local_scores=0.48460972, pem=-0.28448334,
similarity=20.098497]
2: 4212649 Denny_Neagle
    [pem_pos=0, final_scores=0.7374854,
final_local_scores=0.78111064, pem=0.0, similarity=27.399612]
3: 3138 Atlanta
    [pem_pos=0, final_scores=0.43679142,
final_local_scores=0.5164744, pem=-0.23152597,
similarity=20.761415]
4: 2140 Atlanta_Braves
    [pem_pos=0, final_scores=0.44723684,
final_local_scores=0.37988442, pem=-0.06390974,
similarity=15.773774]
5: 4212649 Denny_Neagle
    [pem_pos=0, final_scores=0.6545889,
final_local_scores=0.6782476, pem=0.0, similarity=24.329973]
6: 64290 National_League
    [pem_pos=0, final_scores=0.44706243,
final_local_scores=0.4045087, pem=-0.07431644,
similarity=16.565205]
7: 24270 Pittsburgh_Pirates
    [pem_pos=0, final_scores=0.44689,
final_local_scores=0.3837577, pem=-0.15187988,
similarity=16.367754]
8: 24270 Pittsburgh_Pirates
    [pem_pos=1, final_scores=0.54898304,
final_local_scores=0.4952789, pem=-1.4743205,
similarity=26.88741]
9: 12142087 Ron_Wright_(baseball)
    [pem_pos=1, final_scores=0.46173376,
final_local_scores=0.57053256, pem=-0.25910872,
similarity=22.524616]
10: 1655051 Eugene_Emeralds
    [pem_pos=16, final_scores=0.72160107,
final_local_scores=0.77741385, pem=-5.4119554,
similarity=56.72035]
11: 2140 Atlanta_Braves
    [pem_pos=0, final_scores=0.60641956,
final_local_scores=0.5537674, pem=-0.28448334,
similarity=22.1623]
12: 2456294 National_League_East
    [pem_pos=0, final_scores=0.3719576,
```

```
final_local_scores=0.36514097, pem=0.0, similarity=14.986248]
13: 2140 Atlanta_Braves
    [pem_pos=7, final_scores=0.71400154,
final_local_scores=0.6712844, pem=-4.136474,
similarity=46.616966]
14: 486632 Terry_Pendleton
    [pem_pos=0, final_scores=0.5980374,
final_local_scores=0.6174097, pem=0.0, similarity=22.51445]
15: 5158901 Luis_Polonia
    [pem_pos=0, final_scores=0.609293,
final_local_scores=0.63851726, pem=0.0, similarity=23.144342]
16: 875934 Andruw_Jones
    [pem_pos=0, final_scores=0.53160757,
final_local_scores=0.53591174, pem=0.0, similarity=20.082388]
```

Sample: b'823_SOCCER'

```
SOCCER - [1 1 DUTCH 1] FIRST DIVISION SUMMARY . [2 2 AMSTERDAM 2] 1996-08-28 Summary of [3 3 Dutch 3] first division soccer match played on Thursday : [4 4 NAC Breda 4] 1 ( [5 5 Abdellaoui 5] 20th penalty ) [6 6 NEC Nijmegen 6] 1 ( [7 7 Maurice Graef 7] 36th ) . Halftime 1-1 . Attendance 10,760 .
1: 21148 Netherlands
2: 844 Amsterdam
3: 21148 Netherlands
4: 505173 NAC_Breda
5: 19716558 Yassine_Abdellaoui
6: 834154 N.E.C._(football_club)
7: 9560139 Maurice_Graef
1: 21148 Netherlands
    [pem_pos=0, final_scores=0.65217,
final_local_scores=0.6670403, pem=-0.7943452,
similarity=28.315298]
2: 844 Amsterdam
    [pem_pos=0, final_scores=0.604675,
final_local_scores=0.6312664, pem=-0.0642803,
similarity=23.277529]
3: 21148 Netherlands
    [pem_pos=0, final_scores=0.5035458,
final_local_scores=0.504691, pem=-0.7943452,
similarity=23.470472]
4: 505173 NAC_Breda
    [pem_pos=0, final_scores=0.52696437,
final_local_scores=0.5487544, pem=0.0, similarity=20.46564]
5: 19716558 Yassine_Abdellaoui
    [pem_pos=0, final_scores=0.43750954,
final_local_scores=0.5611822, pem=0.0, similarity=20.83651]
6: 834154 N.E.C._(football_club)
    [pem_pos=0, final_scores=0.6073037,
final_local_scores=0.62438977, pem=0.0, similarity=22.72275]
7: 9560139 Maurice_Graef
    [pem_pos=0, final_scores=0.5143737,
final_local_scores=0.63640684, pem=0.0, similarity=23.081362]
```

Sample: b'824_SOCCER'

SOCCER - [1 1 DUTCH 1] FIRST DIVISION RESULTS / STANDINGS . [2 2 AMSTERDAM 2] 1996-08-29 Result of a [3 3 Dutch 3] first division soccer match played on Thursday : [4 4 NAC Breda 4] 1 [5 5 NEC Nijmegen 5] 1 Played on Wednesday : [6 6 Vitesse Arnhem 6] 1 [7 7 Sparta Rotterdam 7] 1 [8 8 Utrecht 8] 0 [9 9 Twente Enschede 9] 0 [10 10 Groningen 10] 1 [11 11 Roda JC Kerkrade 11] 1 [12 12 Feyenoord 12] 2 Graafschap Doetinchem 1 [13 13 Willem II Tilburg 13] 1 [14 14 RKC Waalwijk 14] 3 [15 15 Volendam 15] 1 [16 16 PSV Eindhoven 16] 3 [17 17 Ajax Amsterdam 17] 1 [18 18 AZ Alkmaar 18] 0 Played on Tuesday : [19 19 Fortuna Sittard 19] 2 [20 20 Heerenveen 20] 4 Standings (tabulate under played , won , drawn , lost , goals for , goals against , points) : [21 21 PSV Eindhoven 21] 3 3 0 0 11 3 9 [22 22 Feyenoord Rotterdam 22] 3 2 1 0 6 2 7 [23 23 Vitesse Arnhem 23] 3 2 1 0 4 1 7 [24 24 Heerenveen 24] 3 2 0 1 7 5 6 [25 25 Ajax Amsterdam 25] 3 2 0 1 2 2 6 [26 26 Twente Enschede 26] 3 1 2 0 4 2 5 [27 27 RKC Waalwijk 27] 3 1 1 1 7 6 4 Graafschap Doetinchem 3 1 1 1 5 5 4 [28 28 NAC Breda 28] 3 1 1 1 2 2 4 [29 29 Fortuna Sittard 29] 3 1 1 1 3 4 4 [30 30 Roda JC Kerkrade 30] 3 0 3 0 3 3 3 [31 31 Utrecht 31] 3 0 2 1 2 3 2 [32 32 Sparta Rotterdam 32] 3 0 2 1 1 2 2 [33 33 Groningen 33] 3 0 2 1 2 5 2 [34 34 NEC Nijmegen 34] 3 0 2 1 2 5 2 [35 35 Willem II Tilburg 35] 3 0 1 2 1 4 1 [36 36 AZ Alkmaar 36] 3 0 1 1 0 3 1 [37 37 Volendam 37] 3 0 1 2 2 7 1
1: 21148 Netherlands
2: 844 Amsterdam
3: 21148 Netherlands
4: 505173 NAC_Breda
5: 834154 N.E.C._(football_club)
6: 834256 Vitesse
7: 79818 Sparta_Rotterdam
8: 834226 FC_Utrecht
9: 800033 FC_Twente
10: 1103112 FC_Groningen
11: 834196 Roda_JC_Kerkrade
12: 47707 Feyenoord
13: 834275 Willem_II_(football_club)
14: 1107824 RKC_Waalwijk
15: 2202202 FC_Volendam
16: 79820 PSV_Eindhoven
17: 2273 AFC_Ajax
18: 2289 AZ_Alkmaar
19: 1336325 Fortuna_Sittard
20: 745757 SC_Heerenveen
21: 79820 PSV_Eindhoven
22: 47707 Feyenoord
23: 834256 Vitesse
24: 745757 SC_Heerenveen
25: 2273 AFC_Ajax
26: 800033 FC_Twente

```
27: 1107824 RKC_Waalwijk
28: 505173 NAC_Breda
29: 1336325 Fortuna_Sittard
30: 834196 Roda_JC_Kerkrade
31: 834226 FC_Utrecht
32: 79818 Sparta_Rotterdam
33: 1103112 FC_Groningen
34: 834154 N.E.C._(football_club)
35: 834275 Willem_II_(football_club)
36: 2289 AZ_Alkmaar
37: 2202202 FC_Volendam
1: 21148 Netherlands
    [pem_pos=0, final_scores=0.6436288,
final_local_scores=0.70943534, pem=-0.7943452,
similarity=29.580452]
2: 844 Amsterdam
    [pem_pos=0, final_scores=0.5660781,
final_local_scores=0.6323258, pem=-0.0642803,
similarity=23.309143]
3: 21148 Netherlands
    [pem_pos=0, final_scores=0.61478746,
final_local_scores=0.6779306, pem=-0.7943452,
similarity=28.640287]
4: 505173 NAC_Breda
    [pem_pos=0, final_scores=0.6201207,
final_local_scores=0.5804351, pem=0.0, similarity=21.411055]
5: 834154 N.E.C._(football_club)
    [pem_pos=0, final_scores=0.60530657,
final_local_scores=0.56007063, pem=0.0, similarity=20.803337]
6: 834256 Vitesse
    [pem_pos=0, final_scores=0.64275223,
final_local_scores=0.60358894, pem=0.0, similarity=22.10201]
7: 79818 Sparta_Rotterdam
    [pem_pos=0, final_scores=0.5695582,
final_local_scores=0.52964306, pem=0.0, similarity=19.895319]
8: 834226 FC_Utrecht
    [pem_pos=2, final_scores=0.84275633,
final_local_scores=0.8218588, pem=-2.2398934,
similarity=40.79651]
9: 800033 FC_Twente
    [pem_pos=0, final_scores=0.5682688,
final_local_scores=0.5171051, pem=0.0, similarity=19.52116]
10: 1103112 FC_Groningen
    [pem_pos=2, final_scores=0.7764453,
final_local_scores=0.74324065, pem=-1.9677447,
similarity=36.970398]
11: 834196 Roda_JC_Kerkrade
    [pem_pos=0, final_scores=0.6294903,
final_local_scores=0.59682834, pem=0.0, similarity=21.90026]
12: 47707 Feyenoord
    [pem_pos=0, final_scores=0.60450596,
final_local_scores=0.5537161, pem=0.0, similarity=20.613707]
```

```
13: 834275 Willem_II_(football_club)
    [pem_pos=0, final_scores=0.77362305,
final_local_scores=0.74201214, pem=0.0, similarity=26.232836]
14: 1107824 RKC_Waalwijk
    [pem_pos=0, final_scores=0.58884436,
final_local_scores=0.5493709, pem=0.0, similarity=20.484035]
15: 2202202 FC_Volendam
    [pem_pos=1, final_scores=0.64351076,
final_local_scores=0.6174086, pem=-0.8958714,
similarity=27.386305]
16: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.52221817,
final_local_scores=0.46432555, pem=0.0, similarity=17.946114]
17: 2273 AFC_Ajax
    [pem_pos=0, final_scores=0.68345594,
final_local_scores=0.6458605, pem=0.0, similarity=23.363478]
18: 2289 AZ_Alkmaar
    [pem_pos=0, final_scores=0.48801523,
final_local_scores=0.42720026, pem=0.0, similarity=16.83822]
19: 1336325 Fortuna_Sittard
    [pem_pos=0, final_scores=0.7512589,
final_local_scores=0.72915995, pem=0.0, similarity=25.8493]
20: 745757 SC_Heerenveen
    [pem_pos=1, final_scores=0.69191283,
final_local_scores=0.6548621, pem=-0.6351416,
similarity=27.086103]
21: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.5080182,
final_local_scores=0.44881427, pem=0.0, similarity=17.483227]
22: 47707 Feyenoord
    [pem_pos=0, final_scores=0.56770223,
final_local_scores=0.5135136, pem=0.0, similarity=19.413984]
23: 834256 Vitesse
    [pem_pos=0, final_scores=0.5790001,
final_local_scores=0.53394943, pem=0.0, similarity=20.023829]
24: 745757 SC_Heerenveen
    [pem_pos=1, final_scores=0.63430876,
final_local_scores=0.59193844, pem=-0.6351416,
similarity=25.208336]
25: 2273 AFC_Ajax
    [pem_pos=0, final_scores=0.5979967,
final_local_scores=0.5525093, pem=0.0, similarity=20.577692]
26: 800033 FC_Twente
    [pem_pos=0, final_scores=0.49726027,
final_local_scores=0.43953907, pem=0.0, similarity=17.206436]
27: 1107824 RKC_Waalwijk
    [pem_pos=0, final_scores=0.54940206,
final_local_scores=0.50628614, pem=0.0, similarity=19.198301]
28: 505173 NAC_Breda
    [pem_pos=0, final_scores=0.69542545,
final_local_scores=0.66269404, pem=0.0, similarity=23.865826]
29: 1336325 Fortuna_Sittard
```

```
[pem_pos=0, final_scores=0.7274866,
final_local_scores=0.7031923, pem=0.0, similarity=25.074373]
30: 834196 Roda_JC_Kerkrade
    [pem_pos=0, final_scores=0.6147418,
final_local_scores=0.58071786, pem=0.0, similarity=21.419493]
31: 834226 FC_Utrecht
    [pem_pos=2, final_scores=0.6795555,
final_local_scores=0.6435866, pem=-2.2398934,
similarity=35.47651]
32: 79818 Sparta_Rotterdam
    [pem_pos=0, final_scores=0.5977387,
final_local_scores=0.5604259, pem=0.0, similarity=20.813938]
33: 1103112 FC_Groningen
    [pem_pos=2, final_scores=0.74956375,
final_local_scores=0.71387655, pem=-1.9677447,
similarity=36.094116]
34: 834154 N.E.C._(football_club)
    [pem_pos=0, final_scores=0.6371175,
final_local_scores=0.5948193, pem=0.0, similarity=21.840307]
35: 834275 Willem_II_(football_club)
    [pem_pos=0, final_scores=0.74018496,
final_local_scores=0.7054862, pem=0.0, similarity=25.142828]
36: 2289 AZ_Alkmaar
    [pem_pos=0, final_scores=0.48095268,
final_local_scores=0.41948557, pem=0.0, similarity=16.608]
37: 2202202 FC_Volendam
    [pem_pos=1, final_scores=0.6693354,
final_local_scores=0.6456181, pem=-0.8958714,
similarity=28.228132]
```

Sample: b'825_ATHLETICS'

ATHLETICS - JOHNSON , [1 1 CHRISTIE 1] , [2 2 BAILEY 2] TO
RUN [3 3 OWENS 3] RELAY . Adrian Warner [4 4 BERLIN 4]
1996-08-29 1 Olympic champions [5 5 Michael Johnson 5] and [6
6 Donovan Bailey 6] and former champion [7 7 Linford Christie
7] will run in a " Dream Team " sprint relay squad in honour
of [8 8 Jesse Owens 8] on Friday . Johnson , the 200 metres
world record holder , and [9 9 Britain 9] 's 1992 2 Olympic
100 champion [10 10 Christie 10] confirmed on Thursday that
they would join 100 record holder [11 11 Bailey 11] and [12 12
Namibia 12] 's [13 13 Frankie Fredericks 13] in one of the
greatest 4x100 squads ever assembled . They will run against
quartets from the [14 14 United States 14] , [15 15 Europe 15]
and [16 16 Africa 16] in a special race at the [17 17 Berlin 1
7] Grand Prix meeting to mark the 60th anniversary of [18 18 0
wens 18] 's four gold medals at the 1936 3 Olympics in the [19
19 German 19] capital . [20 20 Christie 20] will run the
anchor leg after [21 21 Canada 21] 's [22 22 Bailey 22] , [23
23 American 23] Johnson and 4 Olympic silver medallist [24 24
Fredericks 24] have run the first three stages of the relay .
The participation of [25 25 Bailey 25] and [26 26 Fredericks 2
6] had been known before Thursday . But [27 27 Christie 27]

did not announce his decision to run until the eve of the meeting when organisers also confirmed Johnson would take part . [28 28 Christie 28] is due to retire from international competition at the end of the season although he may compete for [29 29 Britain 29] in next season 's [30 30 European Cup 30] . Although [31 31 Christie 31] , who is not racing the individual 100 metres in [32 32 Berlin 32] , took his time to agree to run , the veteran was clearly delighted to be part of the tribute to the black [33 33 American 33] . When five former 5 Olympic 100 champions from 1948 to 1980 , who have been invited to watch the race , turned up at a news conference on Thursday , [34 34 Christie 34] was quick to put his autograph book in front of them . " I do n't normally do this but can you please sign , " he said thrusting an ornate white book in front of [35 35 Americans 35] [36 36 Harrison Dillard 36] (1948) , [37 37 Lindy Remigino 37] (1952) , [38 38 Jim Hines 38] (1968) , [39 39 Trinidad 39] 's [40 40 Hasely Crawford 40] (1976) and [41 41 Britain 41] 's [42 42 Allan Wells 42] (1980) . " [43 43 Jesse Owens 43] inspired everyone here and it is great to have a tribute to him . " [44 44 Owens 44] 's widow Ruth is not well enough to attend but a message from her will be read out by the sprinter 's grand-daughter Gina Tillman during the meeting [45 45 Berlin 45] organisers hoped to have [46 46 American 46] 1984 and 1988 champion [47 47 Carl Lewis 47] in the squad but he injured himself in last Friday 's [48 48 Brussels 48] meeting .

1: 67118 Linford_Christie
2: 67116 Donovan_Bailey
3: 46827 Jesse_Owens
4: 3354 Berlin
5: 653621 Michael_Johnson_(sprinter)
6: 67116 Donovan_Bailey
7: 67118 Linford_Christie
8: 46827 Jesse_Owens
9: 31717 United_Kingdom
10: 67118 Linford_Christie
11: 67116 Donovan_Bailey
12: 21292 Namibia
13: 71410 Frankie_Fredericks
14: 3434750 United_States
15: 9239 Europe
16: 5334607 Africa
17: 3354 Berlin
18: 46827 Jesse_Owens
19: 11867 Germany
20: 67118 Linford_Christie
21: 5042916 Canada
22: 67116 Donovan_Bailey
23: 3434750 United_States
24: 71410 Frankie_Fredericks
25: 67116 Donovan_Bailey

```
26: 71410 Frankie_Fredericks
27: 67118 Linford_Christie
28: 67118 Linford_Christie
29: 31717 United_Kingdom
30: 7108659 European_Cup_(athletics)
31: 67118 Linford_Christie
32: 3354 Berlin
33: 3434750 United_States
34: 67118 Linford_Christie
35: 3434750 United_States
36: 68251 Harrison_Dillard
37: 69112 Lindy_Remigino
38: 67837 Jim_Hines
39: 452912 Trinidad
40: 67506 Hasely_Crawford
41: 31717 United_Kingdom
42: 67299 Allan_Wells
43: 46827 Jesse_Owens
44: 46827 Jesse_Owens
45: 3354 Berlin
46: 3434750 United_States
47: 67127 Carl_Lewis
48: 3708 Brussels
1: 650360 Olympic_sports
    [winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.14458334, final_local_scores=-0.19484156,
pem=-3.2532303, similarity=15.9668255|
     RECALL MISS: 650360 Olympic_sports]
2: 650360 Olympic_sports
    [winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.29617143, final_local_scores=-0.36042848,
pem=-3.2532303, similarity=11.025381|
     RECALL MISS: 650360 Olympic_sports]
3: 39721 1936_Summer_Olympics
    [winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.21901603, final_local_scores=-0.27614796,
pem=-2.7333555, similarity=10.71332|
     RECALL MISS: 39721 1936_Summer_Olympics]
4: 650360 Olympic_sports
    [winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.1941708, final_local_scores=-0.24900833,
pem=-3.2532303, similarity=14.350381|
     RECALL MISS: 650360 Olympic_sports]
5: 650360 Olympic_sports
    [winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.759666, final_local_scores=-0.8667259,
pem=-3.2532303, similarity=-4.0835443|
     RECALL MISS: 650360 Olympic_sports]
1: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.379938,
final_local_scores=1.447315, pem=0.0, similarity=47.28048]
2: 67116 Donovan_Bailey
```

```
[pem_pos=0, final_scores=0.6314735,
final_local_scores=0.6711657, pem=0.0, similarity=24.118637]
3: 46827 Jesse_Owens
    [pem_pos=0, final_scores=1.2184521,
final_local_scores=1.2858546, pem=0.0, similarity=42.46218]
4: 3354 Berlin
    [pem_pos=0, final_scores=0.5442725,
final_local_scores=0.5860868, pem=-0.15483057,
similarity=22.421705]
5: 653621 Michael_Johnson_(sprinter)
    [pem_pos=0, final_scores=0.56048363,
final_local_scores=0.57031155, pem=-0.36135122,
similarity=23.07403]
6: 67116 Donovan_Bailey
    [pem_pos=0, final_scores=0.603215,
final_local_scores=0.6402976, pem=0.0, similarity=23.19747]
7: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.0156438,
final_local_scores=1.046577, pem=0.0, similarity=35.321655]
8: 46827 Jesse_Owens
    [pem_pos=0, final_scores=0.6417747,
final_local_scores=0.6525759, pem=0.0, similarity=23.563879]
9: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.518606,
final_local_scores=0.55185264, pem=-0.61114657,
similarity=23.881607]
10: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.6516078,
final_local_scores=1.7440729, pem=0.0, similarity=56.13633]
11: 67116 Donovan_Bailey
    [pem_pos=0, final_scores=0.817725,
final_local_scores=0.87461716, pem=0.0, similarity=30.190035]
12: 21292 Namibia
    [pem_pos=0, final_scores=0.3698666,
final_local_scores=0.45684022, pem=-0.36777115,
similarity=19.722733]
13: 71410 Frankie_Fredericks
    [pem_pos=0, final_scores=0.8058217,
final_local_scores=0.87658185, pem=0.0, similarity=30.248663]
14: 3434750 United_States
    [pem_pos=0, final_scores=0.52415603,
final_local_scores=0.55968845, pem=-0.073457584,
similarity=21.191406]
15: 9239 Europe
    [pem_pos=0, final_scores=0.61256117,
final_local_scores=0.69788605, pem=-0.086338006,
similarity=25.385546]
16: 5334607 Africa
    [pem_pos=0, final_scores=0.5656236,
final_local_scores=0.6611146, pem=-0.13197383,
similarity=24.536385]
17: 3354 Berlin
```

```
[pem_pos=0, final_scores=0.41481763,
final_local_scores=0.44467705, pem=-0.15483057,
similarity=18.201757]
18: 46827 Jesse_Owens
    [pem_pos=0, final_scores=1.3575667,
final_local_scores=1.4378161, pem=0.0, similarity=46.997017]
19: 11867 Germany
    [pem_pos=1, final_scores=0.78876585,
final_local_scores=0.86472553, pem=-1.0465823,
similarity=35.586327]
20: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.6692493,
final_local_scores=1.7633436, pem=0.0, similarity=56.711403]
21: 5042916 Canada
    [pem_pos=0, final_scores=0.67851895,
final_local_scores=0.7367614, pem=-0.11662793,
similarity=26.710382]
22: 67116 Donovan_Bailey
    [pem_pos=0, final_scores=0.6392511,
final_local_scores=0.6796616, pem=0.0, similarity=24.37217]
23: 3434750 United_States
    [pem_pos=0, final_scores=0.69537824,
final_local_scores=0.7467227, pem=-0.12994486,
similarity=27.080069]
24: 71410 Frankie_Fredericks
    [pem_pos=0, final_scores=1.1728013,
final_local_scores=1.2823241, pem=0.0, similarity=42.356823]
25: 67116 Donovan_Bailey
    [pem_pos=0, final_scores=0.7190562,
final_local_scores=0.7668365, pem=0.0, similarity=26.973644]
26: 71410 Frankie_Fredericks
    [pem_pos=0, final_scores=1.065327,
final_local_scores=1.1649249, pem=0.0, similarity=38.853394]
27: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.6343144,
final_local_scores=1.7251825, pem=0.0, similarity=55.572598]
28: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.5376222,
final_local_scores=1.619561, pem=0.0, similarity=52.42064]
29: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.506202,
final_local_scores=0.5383032, pem=-0.61114657,
similarity=23.477264]
30: 7108659 European_Cup_(athletics)
    [pem_pos=6, final_scores=0.17891131,
final_local_scores=0.2013604, pem=-3.4355228,
similarity=28.781616]
31: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.7232398,
final_local_scores=1.82232, pem=0.0, similarity=58.471375]
32: 3354 Berlin
    [pem_pos=0, final_scores=0.6286784,
```

```
final_local_scores=0.67828745, pem=-0.15483057,
similarity=25.173157]
33: 3434750 United_States
    [pem_pos=0, final_scores=0.4287328,
final_local_scores=0.45545304, pem=-0.12994486,
similarity=18.388]
34: 67118 Linford_Christie
    [pem_pos=0, final_scores=1.5872008,
final_local_scores=1.673718, pem=0.0, similarity=54.036793]
35: 3434750 United_States
    [pem_pos=0, final_scores=0.42157012,
final_local_scores=0.44762892, pem=-0.5581965,
similarity=20.483412]
36: 68251 Harrison_Dillard
    [pem_pos=0, final_scores=0.5549575,
final_local_scores=0.63238424, pem=0.0, similarity=22.96132]
37: 69112 Lindy_Remigino
    [pem_pos=0, final_scores=0.5615357,
final_local_scores=0.6517139, pem=0.0, similarity=23.538157]
38: 67837 Jim_Hines
    [pem_pos=0, final_scores=0.45520264,
final_local_scores=0.50937694, pem=0.0, similarity=19.290537]
39: 452912 Trinidad
    [pem_pos=0, final_scores=0.28748852,
final_local_scores=0.356797, pem=-0.34123632,
similarity=16.592943]
40: 67506 Hasely_Crawford
    [pem_pos=0, final_scores=0.5850908,
final_local_scores=0.6654271, pem=0.0, similarity=23.947384]
41: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5689575,
final_local_scores=0.606854, pem=-0.61114657,
similarity=25.522957]
42: 67299 Allan_Wells
    [pem_pos=0, final_scores=0.63264126,
final_local_scores=0.6792116, pem=0.0, similarity=24.358742]
43: 46827 Jesse_Owens
    [pem_pos=0, final_scores=0.6195151,
final_local_scores=0.6282607, pem=0.0, similarity=22.838264]
44: 46827 Jesse_Owens
    [pem_pos=0, final_scores=1.3148109,
final_local_scores=1.3911119, pem=0.0, similarity=45.603264]
45: 3354 Berlin
    [pem_pos=0, final_scores=0.5699019,
final_local_scores=0.6140831, pem=-0.15483057,
similarity=23.257172]
46: 3434750 United_States
    [pem_pos=0, final_scores=0.51456434,
final_local_scores=0.5492109, pem=-0.12994486,
similarity=21.185923]
47: 67127 Carl_Lewis
    [pem_pos=0, final_scores=0.5928721,
```

```
final_local_scores=0.60854334, pem=0.0, similarity=22.24986]
48: 3708 Brussels
    [pem_pos=0, final_scores=0.48050368,
final_local_scores=0.54517245, pem=-0.025440974,
similarity=20.497099]

Sample: b'826_SOCCER'
SOCCER - [1 1 GENOA 1] AWARDED [2 2 ITALIAN CUP 2] WIN . [3 3
MILAN 3] 1996-08-29 [4 4 Italian 4] soccer 's sports judge on
Thursday ruled that [5 5 Genoa 5] , beaten 3-0 by [6 6 Lecce 6
] in the first round of the [7 7 Italian Cup 7] , should be
awarded a 2-0 victory because their opponents fielded a
banned player . The ruling meant that serie B club [8 8 Genoa
8] now meet local [9 9 serie A 9] arch-rivals [10 10 Sampdoria
10] in the second round . [11 11 Genoa 11] appealed after
their defeat last Saturday on the grounds that [12 12 Lecce 12
] striker [13 13 Jonathan Bachini 13] , who came on in the
71st minute with his team leading 2-0 , still had a one-match
suspension to serve from last season .
1: 30872779 Genoa_C.F.C.
2: 1281200 Coppa_Italia
3: 36511 Milan
4: 14532 Italy
5: 30872779 Genoa_C.F.C.
6: 1135354 U.S._Lecce
7: 1281200 Coppa_Italia
8: 30872779 Genoa_C.F.C.
9: 15496 Serie_A
10: 574307 U.C._Sampdoria
11: 30872779 Genoa_C.F.C.
12: 1135354 U.S._Lecce
13: 2273335 Jonathan_Bachini
1: 30872779 Genoa_C.F.C.
    [pem_pos=1, final_scores=0.66850144,
final_local_scores=0.62154764, pem=-1.6777803,
similarity=31.761965]
2: 1281200 Coppa_Italia
    [pem_pos=0, final_scores=0.5181426,
final_local_scores=0.48497635, pem=0.0, similarity=18.562374]
3: 36511 Milan
    [pem_pos=0, final_scores=0.42260057,
final_local_scores=0.4621572, pem=-0.3015942,
similarity=19.521523]
4: 14532 Italy
    [pem_pos=0, final_scores=0.6146599,
final_local_scores=0.6916359, pem=-0.75948656,
similarity=28.859713]
5: 30872779 Genoa_C.F.C.
    [pem_pos=1, final_scores=0.7078858,
final_local_scores=0.6645691, pem=-1.6777803,
similarity=33.04581]
6: 1135354 U.S._Lecce
```

```
[pem_pos=1, final_scores=0.55140656,
final_local_scores=0.50619805, pem=-0.4613229,
similarity=21.704418]
7: 1281200 Coppa_Italia
    [pem_pos=0, final_scores=0.40900594,
final_local_scores=0.36576116, pem=0.0, similarity=15.004755]
8: 30872779 Genoa_C.F.C.
    [pem_pos=1, final_scores=0.7100089,
final_local_scores=0.6668883, pem=-1.6777803,
similarity=33.11502]
9: 15496 Serie_A
    [pem_pos=0, final_scores=0.5140998,
final_local_scores=0.45626217, pem=-0.1387667,
similarity=18.460121]
10: 574307 U.C._Sampdoria
    [pem_pos=0, final_scores=0.52142435,
final_local_scores=0.46984386, pem=0.0, similarity=18.11079]
11: 30872779 Genoa_C.F.C.
    [pem_pos=1, final_scores=0.63119227,
final_local_scores=0.5807931, pem=-1.6777803,
similarity=30.545767]
12: 1135354 U.S._Lecce
    [pem_pos=1, final_scores=0.46067768,
final_local_scores=0.40709054, pem=-0.4613229,
similarity=18.746853]
13: 2273335 Jonathan_Bachini
    [pem_pos=0, final_scores=0.47384167,
final_local_scores=0.582974, pem=0.0, similarity=21.48682]
```

Sample: b'827_SOCCER'

SOCCER - [1 1 NICE 1] SACK COACH [2 2 EMON 2] . [3 3 NICE 3] , [4 4 France 4] 1996-08-29 Struggling [5 5 French 5] first division side [6 6 Nice 6] on Thursday announced they were parting with coach [7 7 Albert Emon 7] after a string of poor results . Club president Andre Bois said a successor would be named on Friday . A former player for [8 8 Marseille 8] and [9 9 Monaco 9] , [10 10 Emon 10] , 43 , has coached [11 11 Nice 1 1] since 1992 . The announcement came 24 hours after the team from the [12 12 French Riviera 12] lost at home to [13 13 Guingamp 13] 2-1 in a league match . [14 14 Nice 14] are 18th in the table .

```
1: 704478 OGC_Nice
2: 6085283 Albert_Emon
3: 47088 Nice
4: 5843419 France
5: 5843419 France
6: 704478 OGC_Nice
7: 6085283 Albert_Emon
8: 492068 Olympique_de_Marseille
9: 222981 AS_Monaco_FC
10: 6085283 Albert_Emon
11: 704478 OGC_Nice
```

```
12: 266388 French_Riviera
13: 1693070 En_Avant_de_Guingamp
14: 704478 OGC_Nice
1: 704478 OGC_Nice
    [pem_pos=1, final_scores=0.7689573,
final_local_scores=0.74639153, pem=-2.1390686,
similarity=37.996113]
2: 6085283 Albert_Emon
    [pem_pos=0, final_scores=0.6443439,
final_local_scores=0.6822347, pem=0.0, similarity=24.448957]
3: 47088 Nice
    [pem_pos=0, final_scores=0.50256276,
final_local_scores=0.5330005, pem=-0.2595403,
similarity=21.406933]
4: 5843419 France
    [pem_pos=0, final_scores=0.59985036,
final_local_scores=0.64713746, pem=-0.16240834,
similarity=24.284786]
5: 5843419 France
    [pem_pos=0, final_scores=0.7219917,
final_local_scores=0.7805584, pem=-0.80223274,
similarity=31.7458]
6: 704478 OGC_Nice
    [pem_pos=1, final_scores=0.78931886,
final_local_scores=0.7686335, pem=-2.1390686,
similarity=38.65986]
7: 6085283 Albert_Emon
    [pem_pos=0, final_scores=0.701628,
final_local_scores=0.7448089, pem=0.0, similarity=26.316296]
8: 492068 Olympique_de_Marseille
    [pem_pos=1, final_scores=0.64926016,
final_local_scores=0.62398744, pem=-1.4345005,
similarity=30.511778]
9: 222981 AS_Monaco_FC
    [pem_pos=1, final_scores=0.7193884,
final_local_scores=0.70938283, pem=-2.451784,
similarity=38.592293]
10: 6085283 Albert_Emon
    [pem_pos=0, final_scores=0.40210265,
final_local_scores=0.41762298, pem=0.0, similarity=16.552416]
11: 704478 OGC_Nice
    [pem_pos=1, final_scores=0.7489558,
final_local_scores=0.72454286, pem=-2.1390686,
similarity=37.344105]
12: 266388 French_Riviera
    [pem_pos=0, final_scores=0.32627732,
final_local_scores=0.37930346, pem=0.0, similarity=15.408884]
13: 1693070 En_Avant_de_Guingamp
    [pem_pos=0, final_scores=0.478149,
final_local_scores=0.45896304, pem=0.0, similarity=17.786085]
14: 704478 OGC_Nice
    [pem_pos=1, final_scores=0.7728935,
```

```
final_local_scores=0.7506912, pem=-2.1390686,
similarity=38.124424]
```

Sample: b'828_SOCCER'

SOCCER - THREE PULL OUT OF [1 1 DUTCH 1] SQUAD FOR [2 2 BRAZIL 2] . [3 3 AMSTERDAM 3] 1996-08-29 Three [4 4 Dutch 4] players have pulled out of the squad for Saturday 's friendly international soccer match against [5 5 Brazil 5] in [6 6 Amsterdam 6] . [7 7 Ajax 7] defender [8 8 John Veldman 8] and his team mate [9 9 Richard Witschge 9] are injured , while [10 10 PSV 10] midfielder [11 11 Philip Cocu 11] has a fever . [12 12 Dutch 12] coach [13 13 Guus Hiddink 13] called in [14 14 Feyenoord 14] midfielder [15 15 Giovanni van Bronckhorst 15] and [16 16 Vitesse 16] defender Ferdy Vierklau for [17 17 Cocu 17] and [18 18 Veldman 18] , but did not name a replacement for midfielder [19 19 Witschge 19] .

```
1: 9647657 Netherlands_national_football_team
```

```
2: 3383 Brazil
```

```
3: 844 Amsterdam
```

```
4: 9647657 Netherlands_national_football_team
```

```
5: 149286 Brazil_national_football_team
```

```
6: 844 Amsterdam
```

```
7: 2273 AFC_Ajax
```

```
8: 7363720 John_Veldman
```

```
9: 5823946 Richard_Witschge
```

```
10: 79820 PSV_Eindhoven
```

```
11: 1295837 Phillip_Cocu
```

```
12: 9647657 Netherlands_national_football_team
```

```
13: 1345240 Guus_Hiddink
```

```
14: 47707 Feyenoord
```

```
15: 1466668 Giovanni_van_Bronckhorst
```

```
16: 834256 Vitesse
```

```
17: 1295837 Phillip_Cocu
```

```
18: 7363720 John_Veldman
```

```
19: 5823946 Richard_Witschge
```

```
1: 9647657 Netherlands_national_football_team
```

```
    [pem_pos=11, final_scores=0.52420515,
```

```
final_local_scores=0.47933835, pem=-4.675365,
```

```
similarity=43.81949]
```

```
2: 3383 Brazil
```

```
    [pem_pos=0, final_scores=0.43002635,
```

```
final_local_scores=0.5349342, pem=-0.21313342,
```

```
similarity=21.212269]
```

```
3: 844 Amsterdam
```

```
    [pem_pos=0, final_scores=0.5765541,
```

```
final_local_scores=0.6267487, pem=-0.0642803,
```

```
similarity=23.14271]
```

```
4: 9647657 Netherlands_national_football_team
```

```
    [pem_pos=11, final_scores=0.5318762,
```

```
final_local_scores=0.4877178, pem=-4.675365,
```

```
similarity=44.069546]
```

```
5: 149286 Brazil_national_football_team
```

```
[pem_pos=1, final_scores=0.525856,
final_local_scores=0.5186828, pem=-2.5928838,
similarity=33.668747]
6: 844 Amsterdam
    [pem_pos=0, final_scores=0.4466707,
final_local_scores=0.48487073, pem=-0.0642803,
similarity=18.908789]
7: 2273 AFC_Ajax
    [pem_pos=1, final_scores=0.5090981,
final_local_scores=0.46285897, pem=-1.4356393,
similarity=25.709576]
8: 7363720 John_Veldman
    [pem_pos=0, final_scores=0.6786888,
final_local_scores=0.7430294, pem=0.0, similarity=26.263193]
9: 5823946 Richard_Witschge
    [pem_pos=0, final_scores=0.88464147,
final_local_scores=0.9288568, pem=0.0, similarity=31.808651]
10: 79820 PSV_Eindhoven
    [pem_pos=0, final_scores=0.55129224,
final_local_scores=0.51515394, pem=0.0, similarity=19.462934]
11: 1295837 Phillip_Cocu
    [pem_pos=0, final_scores=0.90894514,
final_local_scores=0.9292503, pem=0.0, similarity=31.820395]
12: 9647657 Netherlands_national_football_team
    [pem_pos=11, final_scores=0.50057125,
final_local_scores=0.4535219, pem=-4.675365,
similarity=43.049072]
13: 1345240 Guus_Hiddink
    [pem_pos=0, final_scores=0.7733353,
final_local_scores=0.79450715, pem=0.0, similarity=27.79939]
14: 47707 Feyenoord
    [pem_pos=0, final_scores=0.6164988,
final_local_scores=0.5818401, pem=0.0, similarity=21.452982]
15: 1466668 Giovanni_van_Bronckhorst
    [pem_pos=0, final_scores=0.78104335,
final_local_scores=0.80786145, pem=0.0, similarity=28.19791]
16: 834256 Vitesse
    [pem_pos=0, final_scores=0.70752174,
final_local_scores=0.7114761, pem=0.0, similarity=25.321579]
17: 1295837 Phillip_Cocu
    [pem_pos=0, final_scores=0.810919,
final_local_scores=0.82969403, pem=0.0, similarity=28.849438]
18: 7363720 John_Veldman
    [pem_pos=0, final_scores=0.6345172,
final_local_scores=0.7082349, pem=0.0, similarity=25.224855]
19: 5823946 Richard_Witschge
    [pem_pos=0, final_scores=0.41710418,
final_local_scores=0.42091376, pem=0.0, similarity=16.65062]
```

Sample: b'829_BRITAIN'

[1 1 BRITAIN 1] WELCOMES ROMANIA-HUNGARY TREATY ACCORD . [2 2
BUCHAREST 2] 1996-08-29 [3 3 Britain 3] joined the [4 4 United

States 4] on Thursday in welcoming [5 5 Romania 5] and [6 6 Hungary 6] 's agreement on the text of a much-delayed friendship treaty , which it said would contribute to stability in the area . " The [7 7 United Kingdom 7] believes that such a treaty will contribute positively to the further development of good neighbourly relations between the two countries and enhance the stability of the region , " said the [8 8 British 8] Foreign office statement . The accord , agreed two weeks ago , is expected to end years of disputes over the status of [9 9 Romania 9] 's large ethnic [10 10 Hungarian 10] minority . It will also boost both countries ' ambitions to join [11 11 NATO 11] and the [12 12 European Union 12] . [13 13 Bucharest 13] and [14 14 Budapest 14] say the treaty should be signed in the first half of September . " The [15 15 United Kingdom 15] looks forward to the early signature of the treaty , " the statement said .

1: 31717 United_Kingdom

2: 36877 Bucharest

3: 31717 United_Kingdom

4: 3434750 United_States

5: 25445 Romania

6: 13275 Hungary

7: 31717 United_Kingdom

8: 31717 United_Kingdom

9: 25445 Romania

10: 13275 Hungary

11: 21133 NATO

12: 9317 European_Union

13: 36877 Bucharest

14: 36787 Budapest

15: 31717 United_Kingdom

1: 31717 United_Kingdom

[pem_pos=0, final_scores=0.3164863,
final_local_scores=0.25466183, pem=-0.61114657,
similarity=15.012839]

2: 36877 Bucharest

[pem_pos=0, final_scores=0.5133169,
final_local_scores=0.49351966, pem=-0.01781679,
similarity=18.914215]

3: 31717 United_Kingdom

[pem_pos=0, final_scores=0.54826945,
final_local_scores=0.50784975, pem=-0.61114657,
similarity=22.568474]

4: 3434750 United_States

[pem_pos=0, final_scores=0.49194568,
final_local_scores=0.46120405, pem=-0.073457584,
similarity=18.252436]

5: 25445 Romania

[pem_pos=0, final_scores=0.74116784,
final_local_scores=0.7340363, pem=-0.21241479,
similarity=27.149965]

6: 13275 Hungary

```
[pem_pos=0, final_scores=0.56589115,
final_local_scores=0.5365941, pem=-0.32577527,
similarity=21.874369]
7: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5425297,
final_local_scores=0.50157994, pem=-0.09933331,
similarity=19.59805]
8: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.6185227,
final_local_scores=0.58459073, pem=-0.48936236,
similarity=24.196295]
9: 25445 Romania
    [pem_pos=0, final_scores=0.65201503,
final_local_scores=0.63665044, pem=-0.21241479,
similarity=24.243776]
10: 13275 Hungary
    [pem_pos=1, final_scores=0.6574116,
final_local_scores=0.6365663, pem=-1.0319586,
similarity=28.698072]
11: 21133 NATO
    [pem_pos=0, final_scores=0.43463594,
final_local_scores=0.4670598, pem=0.0, similarity=18.027708]
12: 9317 European_Union
    [pem_pos=0, final_scores=0.49004823,
final_local_scores=0.5000233, pem=-0.005814796,
similarity=19.043028]
13: 36877 Bucharest
    [pem_pos=0, final_scores=0.45014387,
final_local_scores=0.42451268, pem=-0.01781679,
similarity=16.854908]
14: 36787 Budapest
    [pem_pos=0, final_scores=0.5889797,
final_local_scores=0.5829432, pem=-0.030229695,
similarity=21.650293]
15: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.49920034,
final_local_scores=0.4542492, pem=-0.09933331,
similarity=18.185606]
```

Sample: b'830_U.S.'

```
[1 1 U.S. 1] EMBASSY IN [2 2 ATHENS 2] CLOSED ON [3 3 LABOUR
DAY 3] , SEP 2 . [4 4 ATHENS 4] 1996-08-29 The [5 5 U.S. 5]
embassy in [6 6 Athens 6] , the consulates general in [7 7 At
hens 7] and [8 8 Thessaloniki 8] and all [9 9 U.S. 9]
government offices in [10 10 Greece 10] will be closed on
Monday , September 2 in observance of [11 11 Labour Day 11] ,
a [12 12 U.S. 12] national holiday , the embassy said . --
George Georgopoulos , Athens Newsroom +301 3311812-4
```

```
1: 3434750 United_States
2: 1216 Athens
3: 18051 Labour_Day
4: 1216 Athens
```

```
5: 3434750 United_States
6: 1216 Athens
7: 1216 Athens
8: 40471 Thessaloniki
9: 3434750 United_States
10: 12108 Greece
11: 19156272 Labor_Day
12: 3434750 United_States
1: 3434750 United_States
    [pem_pos=0, final_scores=0.49372047,
final_local_scores=0.47531492, pem=-0.14075686,
similarity=19.039516]
2: 1216 Athens
    [pem_pos=0, final_scores=0.44423378,
final_local_scores=0.41010576, pem=-0.28831622,
similarity=17.895996]
3: 18051 Labour_Day
    [pem_pos=0, final_scores=0.34664494,
final_local_scores=0.40074152, pem=0.0, similarity=16.04864]
4: 1216 Athens
    [pem_pos=0, final_scores=0.5079002,
final_local_scores=0.47965163, pem=-0.28831622,
similarity=19.971384]
5: 3434750 United_States
    [pem_pos=0, final_scores=0.5661028,
final_local_scores=0.55438155, pem=-0.14075686,
similarity=21.399021]
6: 1216 Athens
    [pem_pos=0, final_scores=0.49961883,
final_local_scores=0.4706055, pem=-0.28831622,
similarity=19.70143]
7: 1216 Athens
    [pem_pos=0, final_scores=0.50221676,
final_local_scores=0.47344333, pem=-0.28831622,
similarity=19.786116]
8: 40471 Thessaloniki
    [pem_pos=0, final_scores=0.42931908,
final_local_scores=0.41791552, pem=0.0, similarity=16.561146]
9: 3434750 United_States
    [pem_pos=0, final_scores=0.55978495,
final_local_scores=0.5474803, pem=-0.14075686,
similarity=21.193075]
10: 12108 Greece
    [pem_pos=0, final_scores=0.8361518,
final_local_scores=0.8437941, pem=-0.27607638,
similarity=30.771557]
11: 19156272 Labor_Day
    [pem_pos=2, final_scores=0.273947,
final_local_scores=0.31384808, pem=-1.3296652,
similarity=20.686491]
12: 3434750 United_States
    [pem_pos=0, final_scores=0.48724395,
```

```
final_local_scores=0.46824032, pem=-0.14075686,  
similarity=18.828396]
```

Sample: b'831_ND'

ND PARTY PICKS SPOT THOMSON , BOLD / OGILVY , MATHER FOR AD CAMPAIGN . [1 1 ATHENS 1] 1996-08-29 [2 2 Greek 2] conservative New democracy party picked Bold / Ogilvy and Mather advertising companies for its pre-election campaign and Spot Thomson to help party president [3 3 Miltiadis Evert 3] on communication strategy , it said in a statement . Spot Thomson will also be responsible for the campaign TV and radio spots , it said . -- Dimitris Kontogiannis , Athens Newsroom +301 3311812-4

1: 1216 Athens

2: 12108 Greece

3: 5481658 Miltiadis_Evert

1: 1216 Athens

[pem_pos=0, final_scores=0.4665535,
final_local_scores=0.48146307, pem=-0.28831622,
similarity=20.02544]

2: 12108 Greece

[pem_pos=1, final_scores=0.7704883,
final_local_scores=0.8142873, pem=-2.3247876,
similarity=41.032227]

3: 5481658 Miltiadis_Evert

[pem_pos=0, final_scores=0.537353,
final_local_scores=0.6794738, pem=0.0, similarity=24.366568]

Sample: b'832_Rugby'

Rugby star once linked to Princess [1 1 Diana 1] divorces . [2 2 LONDON 2] 1996-08-29 Former [3 3 England 3] rugby captain [4 4 Will Carling 4] , whose marriage broke down after he was romantically linked to Princess [5 5 Diana 5] , was divorced by his wife on Thursday , just 24 hours after [6 6 Diana 6] and [7 7 Prince Charles 7] divorced . [8 8 The Times 8] newspaper said [9 9 Diana 9] was not named in the divorce petition heard by a court in [10 10 Surrey 10] , southern [11 11 England 11] . But [12 12 Julia Carling 12] , a television presenter , is said to have blamed [13 13 Diana 13] for the problems in her marriage and she has repeatedly mocked the princess on her breakfast television programme . [14 14 Diana 14] met [15 15 Will Carling 15] at an exclusive gymnasium in [16 16 London 16] . He has always insisted that they were just friends .

1: 24095 Diana,_Princess_of_Wales

2: 17867 London

3: 9316 England

4: 531810 Will_Carling

5: 24095 Diana,_Princess_of_Wales

6: 24095 Diana,_Princess_of_Wales

7: 125248 Charles,_Prince_of_Wales

8: 39127 The_Times

```
9: 24095 Diana,_Princess_of_Wales
10: 48189 Surrey
11: 9316 England
12: 1485260 Julia_Carling
13: 24095 Diana,_Princess_of_Wales
14: 24095 Diana,_Princess_of_Wales
15: 531810 Will_Carling
16: 17867 London
1: 24095 Diana,_Princess_of_Wales
    [pem_pos=0, final_scores=1.0732771,
final_local_scores=1.0625658, pem=0.0, similarity=35.798798]
2: 17867 London
    [pem_pos=0, final_scores=0.69970435,
final_local_scores=0.7200598, pem=-0.09086737,
similarity=26.071884]
3: 9316 England
    [pem_pos=0, final_scores=0.33522063,
final_local_scores=0.33390296, pem=-0.19053856,
similarity=15.090221]
4: 531810 Will_Carling
    [pem_pos=0, final_scores=0.5046267,
final_local_scores=0.58204615, pem=0.0, similarity=21.459131]
5: 24095 Diana,_Princess_of_Wales
    [pem_pos=0, final_scores=1.0871674,
final_local_scores=1.0777388, pem=0.0, similarity=36.25159]
6: 24095 Diana,_Princess_of_Wales
    [pem_pos=0, final_scores=1.0707617,
final_local_scores=1.059818, pem=0.0, similarity=35.716797]
7: 125248 Charles,_Prince_of_Wales
    [pem_pos=0, final_scores=0.63720727,
final_local_scores=0.6124293, pem=-0.08443151,
similarity=22.824978]
8: 39127 The_Times
    [pem_pos=0, final_scores=0.410366,
final_local_scores=0.41295892, pem=0.0, similarity=16.41323]
9: 24095 Diana,_Princess_of_Wales
    [pem_pos=0, final_scores=1.1425499,
final_local_scores=1.1382358, pem=0.0, similarity=38.05694]
10: 48189 Surrey
    [pem_pos=0, final_scores=0.32802743,
final_local_scores=0.32313123, pem=-0.28536296,
similarity=15.284441]
11: 9316 England
    [pem_pos=0, final_scores=0.5529652,
final_local_scores=0.5717559, pem=-0.19053856,
similarity=22.188227]
12: 1485260 Julia_Carling
    [pem_pos=0, final_scores=0.47088146,
final_local_scores=0.6072122, pem=0.0, similarity=22.210136]
13: 24095 Diana,_Princess_of_Wales
    [pem_pos=0, final_scores=1.1846747,
final_local_scores=1.1842508, pem=0.0, similarity=39.430122]
```

```
14: 24095 Diana,_Princess_of_Wales
    [pem_pos=0, final_scores=1.146878,
final_local_scores=1.1429636, pem=0.0, similarity=38.19803]
15: 531810 Will_Carling
    [pem_pos=0, final_scores=0.4271803,
final_local_scores=0.49744767, pem=0.0, similarity=18.934544]
16: 17867 London
    [pem_pos=0, final_scores=0.505452,
final_local_scores=0.5078686, pem=-0.09086737,
similarity=19.739676]
```

Sample: b'833_London'

```
[1 1 London 1] LIFFE futures APT closing prices . [2 2 LONDON
2] 1996-08-29 London International Financial Futures Exchange
automated pit trading ( APT ) tabular details : CONTRACT
( MONTH ) APT CLOSE SETTLEMENT PREVIOUS SETTLE LONG GILT
( SEP ) ( 1/32 ) 107-12 107-10 107-06 SHORT STERLING ( SEP )
94.26 94.26 94.26 [3 3 GERMAN 3] GOVT BOND ( SEP ) 97.42
97.38 97.34 EUROMARK ( SEP ) 96.84 96.83 96.83 [4 4 ITALIAN 4]
GOVT BOND ( SEP ) 115.62 115.58 115.32 EUROLIRA ( SEP ) 91.37
91.36 91.33 EUROSWISS ( SEP ) 97.79 97.80 97.82 [5 5 FTSE 100
5] ( SEP ) 3,894.00 3,894.00 3,941.50
1: 17867 London
2: 17867 London
3: 11867 Germany
4: 14532 Italy
5: 497717 FTSE_100_Index
1: 17867 London
    [pem_pos=0, final_scores=0.5332078,
final_local_scores=0.49587107, pem=-0.09086737,
similarity=19.381645]
2: 17867 London
    [pem_pos=0, final_scores=0.7313622,
final_local_scores=0.71232474, pem=-0.09086737,
similarity=25.841055]
3: 11867 Germany
    [pem_pos=1, final_scores=0.7137507,
final_local_scores=0.7366093, pem=-1.0465823,
similarity=31.76308]
4: 14532 Italy
    [pem_pos=0, final_scores=0.56996435,
final_local_scores=0.60256463, pem=-0.75948656,
similarity=26.201649]
5: 497717 FTSE_100_Index
    [pem_pos=0, final_scores=0.32137895,
final_local_scores=0.42769402, pem=0.0, similarity=16.852955]
```

Sample: b'834_Zaire'

```
[1 1 Zaire 1] installs first election delegates . [2 2 KINSH
ASA 2] 1996-08-29 A total of 116 delegates to [3 3 Zaire 3]
's National Election Commission ( CNE ) were formally
installed on Thursday , launching another phase of the
```

Central African nation 's much-delayed democratic transition . The 116 , representing political parties in the capital [4 4 Kinshasa 4] , will help organise a voter census , a constitutional referendum planned for January and efforts to brief potential voters on what balloting involves . A total of 9,446 delegates will be deployed throughout [5 5 Zaire 5] 's 11 provinces for the elections , which must be held by July 1997 under the transitional constitution . Presidential , parliamentary and municipal elections are planned for May . " We can meet the required deadlines for organising the elections . All that is needed is for everyone to show goodwill , " [1 Commission 1] spokesman Yoka Lye Mudaba told reporters . Delegates to the commission from the other 10 provinces will be installed progressively from next week with the provinces of 1 North and [6 6 South Kivu 6] , [7 7 Maniema 7] , Shaba and [8 8 Bandundu 8] having priority , he said . The installation of delegates was initially scheduled for July . Officials said lack of funding had delayed the process . President [9 9 Mobutu Sese Seko 9] , who seized power in a 1965 coup , introduced a multi-party system in 1990 but [10 10 Zaire 10] 's transition has lagged well behind that of other states in the region . In the past , [11 11 Mobutu 11] has been elected without opposition .

1: 76762 Democratic_Republic_of_the_Congo
2: 45773 Kinshasa
3: 76762 Democratic_Republic_of_the_Congo
4: 45773 Kinshasa
5: 76762 Democratic_Republic_of_the_Congo
6: 446959 South_Kivu
7: 446957 Maniema
8: 1917423 Bandundu_(city)
9: 64240 Mobutu_Sese_Seko
10: 76762 Democratic_Republic_of_the_Congo
11: 64240 Mobutu_Sese_Seko
1: 446958 North_Kivu
[winner: 21255 North_Korea, pem_pos=2,
final_scores=-0.30015373, final_local_scores=-0.32640293,
pem=-2.7172027, similarity=9.125769]
RECALL MISS: 446958 North_Kivu
1: 76762 Democratic_Republic_of_the_Congo
[pem_pos=3, final_scores=0.38504535,
final_local_scores=0.34068933, pem=-2.0037596,
similarity=25.153318]
2: 45773 Kinshasa
[pem_pos=0, final_scores=0.4345613,
final_local_scores=0.40122938, pem=0.0, similarity=16.063198]
3: 76762 Democratic_Republic_of_the_Congo
[pem_pos=3, final_scores=0.39655298,
final_local_scores=0.35325965, pem=-2.0037596,
similarity=25.528442]
4: 45773 Kinshasa

```
[pem_pos=0, final_scores=0.45253056,
final_local_scores=0.4208581, pem=0.0, similarity=16.648958]
5: 76762 Democratic_Republic_of_the_Congo
    [pem_pos=3, final_scores=0.35996467,
final_local_scores=0.31329247, pem=-2.0037596,
similarity=24.335741]
6: 446959 South_Kivu
    [pem_pos=0, final_scores=0.51277286,
final_local_scores=0.52997506, pem=0.0, similarity=19.905226]
7: 446957 Maniema
    [pem_pos=0, final_scores=0.30357248,
final_local_scores=0.3165981, pem=0.0, similarity=13.537631]
8: 1917423 Bandundu_(city)
    [pem_pos=1, final_scores=0.25315225,
final_local_scores=0.30186567, pem=-1.9972034,
similarity=23.95909]
9: 64240 Mobutu_Sese_Seko
    [pem_pos=0, final_scores=0.56768554,
final_local_scores=0.5859166, pem=0.0, similarity=21.574633]
10: 76762 Democratic_Republic_of_the_Congo
    [pem_pos=3, final_scores=0.4075361,
final_local_scores=0.36525708, pem=-2.0037596,
similarity=25.88647]
11: 64240 Mobutu_Sese_Seko
    [pem_pos=0, final_scores=0.3090034,
final_local_scores=0.30334565, pem=0.0, similarity=13.142152]
1: 9974 European_Commission
    [pem_pos=0, final_scores=0.03329067,
final_local_scores=0.10139386, pem=-0.76247066,
similarity=11.261942]
```

Sample: b'835_Nigeria'

[1 1 Nigeria 1] would not refuse [2 2 Commonwealth 2] officials . [3 3 LAGOS 3] 1996-08-29 [4 4 Nigeria 4] would not object to a visit by [5 5 Commonwealth 5] officials but insists its suspension from the organisation be resolved before any other questions are addressed , Foreign Minister [6 6 Tom Ikimi 6] said on Thursday . [7 7 Ikimi 7] reiterated his position that the [8 8 Commonwealth 8] had no mandate to send a fact-finding mission . " The request I have received is for their officials to come and talk to my officials . We cannot object to people wanting to visit [9 9 Nigeria 9] , " [10 10 Ikimi 10] told [11 11 Reuters 11] by telephone from the capital [12 12 Abuja 12] . " The fundamental problem we have with the [13 13 Commonwealth 13] is our unfair suspension . Any discussions we have at ministerial level will be a continuation of what we began in [14 14 London 14] on that . Before that is accomplished we cannot address anything else . " [15 15 Nigeria 15] was suspended from the [16 16 Commonwealth 16] in November after the execution of [17 17 Ken Saro-Wiwa 17] and eight other minority rights activists in defiance of international pleas for clemency . A meeting of [18 18 Common

wealth 18] ministers in [19 19 London 19] on Wednesday said it planned to send a team of senior officials to [20 20 Nigeria 20] as soon as possible to persuade [21 21 Abuja 21] to accept a fact-finding mission . The timing of that mission has yet to be determined . The latest diplomatic row between [22 22 Nigeria 22] 's military government and the club of [23 2 3 Britain 23] and its former colonies erupted over the terms of a visit by [24 24 Commonwealth 24] ministers to discuss [25 25 Nigeria 25] 's suspension . [26 26 Nigeria 26] said they would be restricted to a two-day meeting with government officials , but the [27 27 Commonwealth 27] said it wanted to hold meetings with people outside the government and called off the visit .

1: 21383 Nigeria
2: 21175158 Commonwealth_of_Nations
3: 85232 Lagos
4: 21383 Nigeria
5: 21175158 Commonwealth_of_Nations
6: 1190018 Tom_Ikimi
7: 1190018 Tom_Ikimi
8: 21175158 Commonwealth_of_Nations
9: 21383 Nigeria
10: 1190018 Tom_Ikimi
11: 18998750 Reuters
12: 56823 Abuja
13: 21175158 Commonwealth_of_Nations
14: 17867 London
15: 21383 Nigeria
16: 21175158 Commonwealth_of_Nations
17: 330970 Ken_Saro-Wiwa
18: 21175158 Commonwealth_of_Nations
19: 17867 London
20: 21383 Nigeria
21: 56823 Abuja
22: 21383 Nigeria
23: 31717 United_Kingdom
24: 21175158 Commonwealth_of_Nations
25: 21383 Nigeria
26: 21383 Nigeria
27: 21175158 Commonwealth_of_Nations
1: 21383 Nigeria
[pem_pos=0, final_scores=0.53769606,
final_local_scores=0.50736064, pem=-0.31455088,
similarity=20.940943]
2: 21175158 Commonwealth_of_Nations
[pem_pos=0, final_scores=0.5402854,
final_local_scores=0.52014315, pem=-0.69133765,
similarity=23.371424]
3: 85232 Lagos
[pem_pos=0, final_scores=0.47485393,
final_local_scores=0.45659214, pem=-0.20534219,
similarity=18.832016]

```
4: 21383 Nigeria
    [pem_pos=0, final_scores=0.57842475,
final_local_scores=0.55185056, pem=-0.31455088,
similarity=22.268612]
5: 21175158 Commonwealth_of_Nations
    [pem_pos=0, final_scores=0.5098665,
final_local_scores=0.4869151, pem=-0.69133765,
similarity=22.379833]
6: 1190018 Tom_Ikimi
    [pem_pos=0, final_scores=0.61181945,
final_local_scores=0.74603075, pem=0.0, similarity=26.352758]
7: 1190018 Tom_Ikimi
    [pem_pos=0, final_scores=0.44414312,
final_local_scores=0.56286985, pem=0.0, similarity=20.886871]
8: 21175158 Commonwealth_of_Nations
    [pem_pos=0, final_scores=0.4738645,
final_local_scores=0.44758838, pem=-0.69133765,
similarity=21.206245]
9: 21383 Nigeria
    [pem_pos=0, final_scores=0.5464509,
final_local_scores=0.51692396, pem=-0.31455088,
similarity=21.226332]
10: 1190018 Tom_Ikimi
    [pem_pos=0, final_scores=0.3754772,
final_local_scores=0.4878627, pem=0.0, similarity=18.64851]
11: 18998750 Reuters
    [pem_pos=0, final_scores=0.32955366,
final_local_scores=0.37503242, pem=0.0, similarity=15.281429]
12: 56823 Abuja
    [pem_pos=0, final_scores=0.40399623,
final_local_scores=0.39944005, pem=0.0, similarity=16.0098]
13: 21175158 Commonwealth_of_Nations
    [pem_pos=0, final_scores=0.47367328,
final_local_scores=0.4473796, pem=-0.69133765,
similarity=21.200014]
14: 17867 London
    [pem_pos=0, final_scores=0.5393,
final_local_scores=0.54939145, pem=-0.09086737,
similarity=20.9788]
15: 21383 Nigeria
    [pem_pos=0, final_scores=0.59135276,
final_local_scores=0.5659724, pem=-0.31455088,
similarity=22.690035]
16: 21175158 Commonwealth_of_Nations
    [pem_pos=0, final_scores=0.5221865,
final_local_scores=0.5003729, pem=-0.69133765,
similarity=22.781439]
17: 330970 Ken_Saro-Wiwa
    [pem_pos=0, final_scores=0.60914963,
final_local_scores=0.68142265, pem=0.0, similarity=24.424725]
18: 21175158 Commonwealth_of_Nations
    [pem_pos=0, final_scores=0.54829484,
```

```
final_local_scores=0.5288922, pem=-0.69133765,
similarity=23.632513]
19: 17867 London
    [pem_pos=0, final_scores=0.5472761,
final_local_scores=0.55810404, pem=-0.09086737,
similarity=21.2388]
20: 21383 Nigeria
    [pem_pos=0, final_scores=0.6438574,
final_local_scores=0.6233258, pem=-0.31455088,
similarity=24.401575]
21: 56823 Abuja
    [pem_pos=0, final_scores=0.314008,
final_local_scores=0.3011416, pem=0.0, similarity=13.076379]
22: 21383 Nigeria
    [pem_pos=0, final_scores=0.6082273,
final_local_scores=0.5844053, pem=-0.31455088,
similarity=23.24011]
23: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.55309886,
final_local_scores=0.53974575, pem=-0.61114657,
similarity=23.520313]
24: 21175158 Commonwealth_of_Nations
    [pem_pos=0, final_scores=0.5266815,
final_local_scores=0.50528294, pem=-0.69133765,
similarity=22.927965]
25: 21383 Nigeria
    [pem_pos=0, final_scores=0.59586626,
final_local_scores=0.5709027, pem=-0.31455088,
similarity=22.837166]
26: 21383 Nigeria
    [pem_pos=0, final_scores=0.5915155,
final_local_scores=0.5661502, pem=-0.31455088,
similarity=22.695341]
27: 21175158 Commonwealth_of_Nations
    [pem_pos=0, final_scores=0.6235382,
final_local_scores=0.6110842, pem=-0.69133765,
similarity=26.085287]
```

Sample: b'836_Dutch'

[1 1 Dutch 1] Queen [2 2 Beatrix 2] to visit S. Africa in October . [3 3 JOHANNESBURG 3] 1996-08-29 Queen [4 4 Beatrix 4] of [5 5 The Netherlands 5] will pay a four-day state visit to [6 6 South Africa 6] in October , the first by a ruling [7 7 Dutch 7] monarch , the [8 8 South African 8] foreign ministry said on Thursday . She will be accompanied by several officials who will sign a cultural agreement with [9 9 South Africa 9] , where the [10 10 Dutch 10] were the first [1 11 European 11] settlers in 1652 . The queen will be accompanied by her husband Prince Claus on the September 30 to October 3 visit .

1: 21148 Netherlands

2: 80482 Beatrix_of_the_Netherlands

```
3: 58893 Johannesburg
4: 80482 Beatrix_of_the_Netherlands
5: 21148 Netherlands
6: 17416221 South_Africa
7: 21148 Netherlands
8: 17416221 South_Africa
9: 17416221 South_Africa
10: 21148 Netherlands
11: 9239 Europe
1: 21148 Netherlands
    [pem_pos=0, final_scores=0.6296528,
final_local_scores=0.5969896, pem=-0.7943452,
similarity=26.224846]
2: 80482 Beatrix_of_the_Netherlands
    [pem_pos=0, final_scores=0.709429,
final_local_scores=0.75178605, pem=0.0, similarity=26.524508]
3: 58893 Johannesburg
    [pem_pos=0, final_scores=0.51921016,
final_local_scores=0.5262036, pem=-0.09562059,
similarity=20.312677]
4: 80482 Beatrix_of_the_Netherlands
    [pem_pos=0, final_scores=0.66812164,
final_local_scores=0.706664, pem=0.0, similarity=25.177977]
5: 21148 Netherlands
    [pem_pos=0, final_scores=0.6032302,
final_local_scores=0.5681268, pem=-0.10774545,
similarity=21.629686]
6: 17416221 South_Africa
    [pem_pos=0, final_scores=0.4066674,
final_local_scores=0.39572567, pem=-0.26251382,
similarity=17.326548]
7: 21148 Netherlands
    [pem_pos=3, final_scores=0.54165465,
final_local_scores=0.50086486, pem=-0.7943452,
similarity=23.356293]
8: 17416221 South_Africa
    [pem_pos=0, final_scores=0.44702798,
final_local_scores=0.4398135, pem=-0.1389451,
similarity=17.97023]
9: 17416221 South_Africa
    [pem_pos=0, final_scores=0.53860444,
final_local_scores=0.5398469, pem=-0.26251382,
similarity=21.627413]
10: 21148 Netherlands
    [pem_pos=3, final_scores=0.40770155,
final_local_scores=0.35454148, pem=-0.7943452,
similarity=18.98971]
11: 9239 Europe
    [pem_pos=0, final_scores=0.4965139,
final_local_scores=0.51495916, pem=-0.25557294,
similarity=20.846966]
```

Sample: b'837_Senegal'
[1 1 Senegal 1] bans guns ahead of local elections . [2 2 DA
KAR 2] 1996-08-29 With tension rising among [3 3 Senegal 3]
's political parties ahead of local elections on November
24 , the [4 4 interior ministry 4] on Thursday banned the
carrying of guns and ammunition until the end of the year . "
It is forbidden for holders of permits for weapons of all
categories to transport the said arms and their ammunition
outside their homes , " the statement said . It said the ban
applied to [5 5 Senegalese 5] nationals and foreign
residents .

1: 27278 Senegal
2: 44251 Dakar
3: 27278 Senegal
4: 354620 Interior_ministry
5: 27278 Senegal
1: 27278 Senegal
 [pem_pos=0, final_scores=0.5423688,
final_local_scores=0.48927295, pem=-0.40308514,
similarity=20.882633]
2: 44251 Dakar
 [pem_pos=0, final_scores=0.4600317,
final_local_scores=0.4439187, pem=-0.008440792,
similarity=17.383034]
3: 27278 Senegal
 [pem_pos=0, final_scores=0.58960074,
final_local_scores=0.5408666, pem=-0.40308514,
similarity=22.42229]
4: 354620 Interior_ministry
 [pem_pos=0, final_scores=0.4868998,
final_local_scores=0.5962407, pem=-0.57122487,
similarity=24.989134]
5: 27278 Senegal
 [pem_pos=0, final_scores=0.5364888,
final_local_scores=0.4828499, pem=0.0, similarity=18.498917]

Sample: b'838_Chad'
Chad parliamentary election set for November 24 . N'DJAMENA
1996-08-29 Chad 's President [1 1 Idriss Déby 1] has signed a
decree fixing November 24 as the date for parliamentary
elections , state radio said on Thursday . Nomads will vote
at mobile polling stations around the vast Central African
country between November 20 and 24 . The electoral commission
said the new 125-member national assembly would be installed
on February 10 . Déby took power in an armed uprising in
1990 . He won 69 percent of votes in the second round of
presidential elections on July 3 , 1996 .. His supporters
will contest the parliamentary election in a coalition of 30
mainly small parties , the Republican Front , led by Déby 's [
2 2 Patriotic Salvation Movement 2] .

1: 334090 Idriss_Déby
2: 405963 Patriotic_Salvation_Movement

```
1: 334090 Idriss_Déby
    [pem_pos=0, final_scores=0.5770543,
final_local_scores=0.6193197, pem=0.0, similarity=22.571447]
2: 405963 Patriotic_Salvation_Movement
    [pem_pos=0, final_scores=0.5378013,
final_local_scores=0.5764418, pem=0.0, similarity=21.291887]
```

Sample: b'839_PRESS'

PRESS DIGEST - [1 1 Ivory_Coast 1] - Aug 29 . [2 2 ABIDJAN 2]
1996-08-29 These are significant stories in the [3 3 Ivorian 3]
] press on Thursday . [4 4 Reuters 4] has not verified these
stories and does not vouch for their accuracy . FRATERNITE
MATIN - Cabinet meeting establishes five new administrative
regions and four new departments as part of government
decentralisation policy . LA VOIE - Members of parliament
seek higher pay and more benefits . The speaker of
parliament , Charles Bauza Donwahi , will meet President [5 5
Henri Konan Bedie 5] on September 3 to discuss their
request . - Deputy director of animal health department
Douati Alphonse says his agents have seized 46 tonnes of
illicit pork in a two-week operation to ensure compliance
with a ban imposed after an outbreak of swine fever . [6 6 LE
JOUR 6] - Raphael Lakpe , publisher of the daily Le
Populaire , was released on Wednesday evening after three
days in custody and will appear in court on Thursday
morning . - Cabinet meeting appoints Colonel Severin Konan
Kouame as gendarmerie commander , replacing General Joseph
Tanny , who has been appointed secretary-general of the
National Security Council . -- [7 7 Abidjan 7] newsroom +225
21 90 90

```
1: 57744 Ivory_Coast
2: 21371767 Abidjan
3: 57744 Ivory_Coast
4: 18998750 Reuters
5: 418112 Henri_Konan_Bédié
6: 7362858 Le_Jour
7: 21371767 Abidjan
1: 57744 Ivory_Coast
    [pem_pos=0, final_scores=0.5402319,
final_local_scores=0.54181945, pem=0.0, similarity=20.258686]
2: 21371767 Abidjan
    [pem_pos=0, final_scores=0.42410272,
final_local_scores=0.39069635, pem=0.0, similarity=15.748872]
3: 57744 Ivory_Coast
    [pem_pos=0, final_scores=0.46660608,
final_local_scores=0.4613945, pem=0.0, similarity=17.858644]
4: 18998750 Reuters
    [pem_pos=0, final_scores=0.2745524,
final_local_scores=0.3372323, pem=0.0, similarity=14.153398]
5: 418112 Henri_Konan_Bédié
    [pem_pos=0, final_scores=0.62345535,
final_local_scores=0.7216991, pem=0.0, similarity=25.626654]
```

```
6: 7362858 Le_Jour
    [pem_pos=0, final_scores=0.3185374,
final_local_scores=0.4459595, pem=0.0, similarity=17.398035]
7: 21371767 Abidjan
    [pem_pos=0, final_scores=0.33806872,
final_local_scores=0.29671726, pem=0.0, similarity=12.944347]
```

Sample: b'840_NATO'

[1 1 NATO 1] releases [2 2 Serb 2] police , crisis easing - [3 3 NATO 3] . [4 4 SARAJEVO 4] 1996-08-29 [5 5 NATO 5] forces released a group of Bosnian Serb policemen late on Thursday and a tense confrontation appeared to be easing , an alliance spokesman said . " The situation in Mahala seems to be very much on its way toward resolution . My understanding is the ([6 6 Serb 6]) police have been released ... In [7 7 Zvornik 7] , we think the situation is winding down as well , " [8 8 NATO 8] spokesman Lieutenant Colonel Max Marriner told [9 9 Reuters 9] . [10 10 NATO 10] troops detained 65 Bosnian Serb policemen early on Thursday after they attacked [11 11 Moslem 11] refugees returning to homes in Mahala , a Serb-controlled village on Bosnia 's internal boundary line . In apparent retaliation for [12 12 NATO 12] 's detention of the [13 13 Serbs 13] , an angry [14 14 Serb 14] mob including policemen trapped six unarmed [15 15 U.N. 15] police monitors and three local aides in their office in the town of [16 16 Zvornik 16] , east of Mahala . Marriner said [17 17 NATO 17] forces confiscated 25 long-barreled [18 18 AK-47 18] automatic assault rifles from the detained [19 19 Serbs 19] before setting them free . " Local radio and television there (in [20 20 Zvornik 20]) are advising people to step back and take it easy . We think that when the word about how things are going in Mahala which is about 35 minutes away by road reaches [21 21 Zvornik 21] that should help , " Marriner said .

```
1: 21133 NATO
2: 29265 Serbia
3: 21133 NATO
4: 26786 Sarajevo
5: 21133 NATO
6: 29265 Serbia
7: 1714456 Zvornik
8: 21133 NATO
9: 18998750 Reuters
10: 21133 NATO
11: 6037917 Islam
12: 21133 NATO
13: 29265 Serbia
14: 29265 Serbia
15: 31769 United_Nations
16: 1714456 Zvornik
17: 21133 NATO
18: 1348 AK-47
```

19: 29265 Serbia
20: 1714456 Zvornik
21: 1714456 Zvornik
1: 21133 NATO
 [pem_pos=0, final_scores=0.40151417,
final_local_scores=0.37090135, pem=0.0, similarity=15.158149]
2: 29265 Serbia
 [pem_pos=1, final_scores=0.51856863,
final_local_scores=0.50089693, pem=-2.1726613,
similarity=30.852747]
3: 21133 NATO
 [pem_pos=0, final_scores=0.44844013,
final_local_scores=0.4221608, pem=0.0, similarity=16.687834]
4: 26786 Sarajevo
 [pem_pos=0, final_scores=0.420515,
final_local_scores=0.40246034, pem=-0.0017846047,
similarity=16.109636]
5: 21133 NATO
 [pem_pos=0, final_scores=0.4375263,
final_local_scores=0.4102391, pem=0.0, similarity=16.332066]
6: 29265 Serbia
 [pem_pos=1, final_scores=0.46615404,
final_local_scores=0.44364196, pem=-2.1726613,
similarity=29.144144]
7: 1714456 Zvornik
 [pem_pos=0, final_scores=0.39349192,
final_local_scores=0.40896147, pem=0.0, similarity=16.29394]
8: 21133 NATO
 [pem_pos=0, final_scores=0.43550843,
final_local_scores=0.40803492, pem=0.0, similarity=16.266289]
9: 18998750 Reuters
 [pem_pos=0, final_scores=0.358167,
final_local_scores=0.41000378, pem=0.0, similarity=16.325043]
10: 21133 NATO
 [pem_pos=0, final_scores=0.38619024,
final_local_scores=0.3541623, pem=0.0, similarity=14.658623]
11: 6037917 Islam
 [pem_pos=1, final_scores=0.34537572,
final_local_scores=0.4165674, pem=-1.2150909,
similarity=23.128767]
12: 21133 NATO
 [pem_pos=0, final_scores=0.42330974,
final_local_scores=0.3947097, pem=0.0, similarity=15.868638]
13: 29265 Serbia
 [pem_pos=4, final_scores=0.42730606,
final_local_scores=0.40120643, pem=-2.3277044,
similarity=28.720932]
14: 29265 Serbia
 [pem_pos=1, final_scores=0.58016115,
final_local_scores=0.56817746, pem=-2.1726613,
similarity=32.86053]
15: 31769 United_Nations

```
[pem_pos=0, final_scores=0.5066027,
final_local_scores=0.5146779, pem=0.0, similarity=19.448727]
16: 1714456 Zvornik
    [pem_pos=0, final_scores=0.44831777,
final_local_scores=0.46885037, pem=0.0, similarity=18.081142]
17: 21133 NATO
    [pem_pos=0, final_scores=0.38628852,
final_local_scores=0.35426962, pem=0.0, similarity=14.661825]
18: 1348 AK-47
    [pem_pos=0, final_scores=0.50273037,
final_local_scores=0.5985893, pem=0.0, similarity=21.952812]
19: 29265 Serbia
    [pem_pos=4, final_scores=0.4164129,
final_local_scores=0.38930732, pem=-2.3277044,
similarity=28.365837]
20: 1714456 Zvornik
    [pem_pos=0, final_scores=0.47913688,
final_local_scores=0.50251555, pem=0.0, similarity=19.08578]
21: 1714456 Zvornik
    [pem_pos=0, final_scores=0.43815392,
final_local_scores=0.45774794, pem=0.0, similarity=17.749825]
```

Sample: b'841_Storm'

Storm kills 11 at [1 1 Macedonian 1] religious festival . [2 2 SKOPJE 2] 1996-08-29 At least 11 people were killed and 60 others injured on Thursday when lightning struck a group of people attending a religious festival in [3 3 Macedonia 3] , police and municipal officials said . Police in [4 4 Berovo 4] , 150 km (90 miles) west of the capital [5 5 Skopje 5] , said there were around 15,000 people gathered around the town 's cathedral when the lightning struck the group during a thunderstorm .

```
1: 23564616 Republic_of_Macedonia
2: 29618 Skopje
3: 23564616 Republic_of_Macedonia
4: 1435264 Berovo
5: 29618 Skopje
1: 23564616 Republic_of_Macedonia
    [pem_pos=3, final_scores=0.59295183,
final_local_scores=0.5705614, pem=-2.0914776,
similarity=32.490185]
2: 29618 Skopje
    [pem_pos=0, final_scores=0.49498618,
final_local_scores=0.45538604, pem=0.0, similarity=17.67934]
3: 23564616 Republic_of_Macedonia
    [pem_pos=0, final_scores=0.4399615,
final_local_scores=0.40344268, pem=-1.1139666,
similarity=22.18717]
4: 1435264 Berovo
    [pem_pos=0, final_scores=0.3885554,
final_local_scores=0.5152804, pem=0.0, similarity=19.46671]
5: 29618 Skopje
```

```
[pem_pos=0, final_scores=0.48886806,
final_local_scores=0.44870287, pem=0.0, similarity=17.479902]
```

Sample: b'842_Kornblum'

[1 1 Kornblum 1] , [2 2 Milosevic 2] discuss election crisis . Peter Greste [3 3 BELGRADE 3] 1996-08-29 [4 4 U.S. 4] envoy [5 5 John Kornblum 5] met [6 6 Serbian 6] President [7 7 Slobodan Milosevic 7] on Thursday in an effort to defuse a growing crisis surrounding Bosnia 's post-war elections . The [8 8 American 8] diplomat arrived in [9 9 Belgrade 9] two days after international organisers postponed municipal elections in Bosnia due to irregularities in the registration of [10 10 Serb 10] refugee voters . " We discussed the decision to postpone the municipal elections and I made clear it was primarily the manipulation of voter registration by the [11 11 Republika Srpska 11] ([12 12 Bosnian Serb 12] republic) which led to this development , " [13 13 Kornblum 1 3] said after three hours of talks . But he said the [14 14 United States 14] still believed it was important to go ahead with national elections in Bosnia as scheduled on September 14 to bolster the peace process . [15 15 Kornblum 15] gave no indication that he had won any specific commitment from [16 16 Milosevic 16] , the patron of the Bosnian Serbs , to rectify any abuses in the registration process . Bosnia 's [17 17 Mos lem 17] political parties have urged their refugees to boycott the elections until irregularities with voter registration are resolved . Human rights workers say [18 18 Serbian 18] and [19 19 Bosnian Serb 19] authorities coerced refugees to register only in Serb-held territory in Bosnia to solidify the results of wartime expulsions and military conquest . [20 20 Serbian 20] officials have denied any abuses occurred during a 10-day registration period and the Bosnian Serbs , angry at the postponement of municipal elections , have threatened to hold local polls on their territory without the international community 's blessing . [2 1 21 Kornblum 21] said only elections endorsed by the Organisation for Security and Cooperation in Europe (OSCE) would be legitimate . " As we have said publicly before , if there is any effort to do so (hold local elections) these elections will not be valid , " he said . " The elections which are valid are those conducted by the international community under the management of the OSCE . " After meeting [22 22 Milosevic 22] , [23 23 Kornblum 23] flew to the [24 24 Croatian 24] capital [25 25 Zagreb 25] . He was due to head to the Bosnian town of [26 26 Banja Luka 26] on Friday to meet [2 7 27 Bosnian Serb 27] acting president [28 28 Biljana Plavsic 28] and [29 29 Bosnian Serb 29] opposition leaders . He also planned to travel to [30 30 Sarajevo 30] to oversee the formal dissolution of the separatist [31 31 Croat 31] mini-state in western Bosnia . [32 32 U.N. 32] police , relief workers and [33 33 NATO 33] officers have reported a rise in political violence across Bosnia in the run-up to the

September 14 elections . Voters will be choosing a three-member presidency and a parliament to rule over a loose union of Bosnia , comprised of a [34 34 Serb 34] republic and a Moslem-Croat federation .

1: 7600227 John_C._Kornblum
2: 25862556 Slobodan_Milošević
3: 55904 Belgrade
4: 3434750 United_States
5: 7600227 John_C._Kornblum
6: 29265 Serbia
7: 25862556 Slobodan_Milošević
8: 3434750 United_States
9: 55904 Belgrade
10: 29265 Serbia
11: 347422 Republika_Srpska
12: 347422 Republika_Srpska
13: 7600227 John_C._Kornblum
14: 3434750 United_States
15: 7600227 John_C._Kornblum
16: 25862556 Slobodan_Milošević
17: 6037917 Islam
18: 29265 Serbia
19: 347422 Republika_Srpska
20: 29265 Serbia
21: 7600227 John_C._Kornblum
22: 25862556 Slobodan_Milošević
23: 7600227 John_C._Kornblum
24: 5573 Croatia
25: 55906 Zagreb
26: 385109 Banja_Luka
27: 347422 Republika_Srpska
28: 158677 Biljana_Plavšić
29: 347422 Republika_Srpska
30: 26786 Sarajevo
31: 5573 Croatia
32: 31769 United_Nations
33: 21133 NATO
34: 29265 Serbia
1: 7600227 John_C._Kornblum
 [pem_pos=0, final_scores=0.7667343,
final_local_scores=0.81947714, pem=0.0, similarity=28.544546]
2: 25862556 Slobodan_Milošević
 [pem_pos=0, final_scores=0.45497096,
final_local_scores=0.45564467, pem=0.0, similarity=17.68706]
3: 55904 Belgrade
 [pem_pos=0, final_scores=0.45993215,
final_local_scores=0.4337566, pem=-0.05193398,
similarity=17.3163]
4: 3434750 United_States
 [pem_pos=0, final_scores=0.4953124,
final_local_scores=0.5218082, pem=-0.14075686,
similarity=20.426968]

5: 7600227 John_C._Kornblum
[pem_pos=0, final_scores=0.67857367,
final_local_scores=0.71684676, pem=0.0, similarity=25.48185]
6: 29265 Serbia
[pem_pos=1, final_scores=0.5331146,
final_local_scores=0.5185605, pem=-1.2504034,
similarity=26.364483]
7: 25862556 Slobodan_Milošević
[pem_pos=0, final_scores=0.5771238,
final_local_scores=0.5890782, pem=0.0, similarity=21.668982]
8: 3434750 United_States
[pem_pos=0, final_scores=0.48380703,
final_local_scores=0.5092403, pem=-0.12994486,
similarity=19.993122]
9: 55904 Belgrade
[pem_pos=0, final_scores=0.47783673,
final_local_scores=0.45331466, pem=-0.05193398,
similarity=17.899952]
10: 29265 Serbia
[pem_pos=1, final_scores=0.6305125,
final_local_scores=0.6249529, pem=-2.1726613,
similarity=34.554825]
11: 347422 Republika_Srpska
[pem_pos=0, final_scores=0.35526878,
final_local_scores=0.3567703, pem=0.0, similarity=14.73645]
12: 347422 Republika_Srpska
[pem_pos=0, final_scores=0.5088364,
final_local_scores=0.52451956, pem=0.0, similarity=19.742422]
13: 7600227 John_C._Kornblum
[pem_pos=0, final_scores=0.693549,
final_local_scores=0.73953336, pem=0.0, similarity=26.158863]
14: 3434750 United_States
[pem_pos=0, final_scores=0.43802625,
final_local_scores=0.4592318, pem=-0.073457584,
similarity=18.19358]
15: 7600227 John_C._Kornblum
[pem_pos=0, final_scores=0.69718474,
final_local_scores=0.7435049, pem=0.0, similarity=26.277382]
16: 25862556 Slobodan_Milošević
[pem_pos=0, final_scores=0.4881422,
final_local_scores=0.49187917, pem=0.0, similarity=18.768368]
17: 6037917 Islam
[pem_pos=1, final_scores=0.37059557,
final_local_scores=0.45720494, pem=-1.2150909,
similarity=24.341473]
18: 29265 Serbia
[pem_pos=1, final_scores=0.61681813,
final_local_scores=0.60999393, pem=-1.2504034,
similarity=29.093037]
19: 347422 Republika_Srpska
[pem_pos=0, final_scores=0.43259853,
final_local_scores=0.44124132, pem=0.0, similarity=17.257235]

20: 29265 Serbia
[pem_pos=1, final_scores=0.57897073,
final_local_scores=0.5686514, pem=-1.2504034,
similarity=27.859293]
21: 7600227 John_C._Kornblum
[pem_pos=0, final_scores=0.67454046,
final_local_scores=0.7187695, pem=0.0, similarity=25.539228]
22: 25862556 Slobodan_Milošević
[pem_pos=0, final_scores=0.45333487,
final_local_scores=0.45385748, pem=0.0, similarity=17.633726]
23: 7600227 John_C._Kornblum
[pem_pos=0, final_scores=0.672033,
final_local_scores=0.7160304, pem=0.0, similarity=25.457489]
24: 5573 Croatia
[pem_pos=1, final_scores=0.5053573,
final_local_scores=0.4974113, pem=-1.1038951,
similarity=24.936611]
25: 55906 Zagreb
[pem_pos=0, final_scores=0.49441117,
final_local_scores=0.48254287, pem=-0.06669851,
similarity=18.852472]
26: 385109 Banja_Luka
[pem_pos=0, final_scores=0.42202765,
final_local_scores=0.42302436, pem=0.0, similarity=16.713604]
27: 347422 Republika_Srpska
[pem_pos=0, final_scores=0.40288645,
final_local_scores=0.40878534, pem=0.0, similarity=16.288683]
28: 158677 Biljana_Plavšić
[pem_pos=0, final_scores=0.59450346,
final_local_scores=0.6810685, pem=0.0, similarity=24.414154]
29: 347422 Republika_Srpska
[pem_pos=0, final_scores=0.53857696,
final_local_scores=0.55700666, pem=0.0, similarity=20.711903]
30: 26786 Sarajevo
[pem_pos=0, final_scores=0.40271693,
final_local_scores=0.37336373, pem=-0.0017846047,
similarity=15.241337]
31: 5573 Croatia
[pem_pos=1, final_scores=0.29414576,
final_local_scores=0.26669475, pem=-1.4398314,
similarity=19.878443]
32: 31769 United_Nations
[pem_pos=0, final_scores=0.44807595,
final_local_scores=0.4756031, pem=0.0, similarity=18.282658]
33: 21133 NATO
[pem_pos=0, final_scores=0.3834116,
final_local_scores=0.3907234, pem=0.0, similarity=15.749679]
34: 29265 Serbia
[pem_pos=1, final_scores=0.6369918,
final_local_scores=0.6320306, pem=-2.1726613,
similarity=34.766037]

Sample: b'843_Gazprom'
[1 1 Gazprom 1] rises to 2,901.48 rbls at auction . [2 2 MOS
COW 2] 1996-08-29 A stake of 10 million shares in [3 3 Russian
3] gas monopoly [4 4 RAO Gazprom 4] was sold at auction on
Thursday at an average 2,901.48 roubles a share , up from
2,891.00 roubles a week ago , the Federal Securities
Corporation (FFK) said . Starting price at the auction was
2,746 roubles a share and the 40 lots sold for between 2,840
and 2,998 roubles a share , the FFK said in a statement . The
stake represented 0.042 percent of [5 5 Gazprom 5] 's
capital . The FFK said that since auctions began , 139.75
million shares , equivalent to 0.59 percent of [6 6 Gazprom 6]
, have changed hands . [7 7 Gazprom 7] , [8 8 Russia 8] 's
biggest company by market capitalisation , has massive
reserves and potentially huge earnings . However , its 23.6
billion shares are highly illiquid and management permission
is required to sell them . [9 9 Gazprom 9] has recently
tightened these rules , making it harder for shareholders to
sell to whoever they want , when they want . The company has
organised regular auctions of its shares to create an orderly
market in the paper . At the first auction on March 6 ,
shares sold for an average 406.6 roubles each , and prices
have been rising steadily since then , but the rise in price
this week and last was much less than in previous auctions .
On the [10 10 Russian Trading System 10] , [11 11 Gazprom 11]
shares rose 25 percent on Thursday to \$ 0.375 from \$ 0.300 ,
after falling by over a third earlier this week . -- Artyom
Danielyan , Moscow Newsroom , +7095 941 8520

1: 320489 Gazprom
2: 19004 Moscow
3: 25391 Russia
4: 320489 Gazprom
5: 320489 Gazprom
6: 320489 Gazprom
7: 320489 Gazprom
8: 25391 Russia
9: 320489 Gazprom
10: 7191778 Russian_Trading_System
11: 320489 Gazprom
1: 320489 Gazprom
[pem_pos=0, final_scores=0.5034581,
final_local_scores=0.43747133, pem=0.0, similarity=17.14473]
2: 19004 Moscow
[pem_pos=0, final_scores=0.6309771,
final_local_scores=0.68440765, pem=-0.106582455,
similarity=25.093414]
3: 25391 Russia
[pem_pos=1, final_scores=0.743344,
final_local_scores=0.78597987, pem=-1.0515022,
similarity=33.263153]
4: 320489 Gazprom
[pem_pos=0, final_scores=0.5771412,

```
final_local_scores=0.51795894, pem=0.0, similarity=19.54664]
5: 320489 Gazprom
    [pem_pos=0, final_scores=0.5157381,
final_local_scores=0.45088536, pem=0.0, similarity=17.54503]
6: 320489 Gazprom
    [pem_pos=0, final_scores=0.5913884,
final_local_scores=0.5335218, pem=0.0, similarity=20.011066]
7: 320489 Gazprom
    [pem_pos=0, final_scores=0.5948518,
final_local_scores=0.53730506, pem=0.0, similarity=20.123968]
8: 25391 Russia
    [pem_pos=0, final_scores=0.81768376,
final_local_scores=0.8671848, pem=-0.20299491,
similarity=31.072153]
9: 320489 Gazprom
    [pem_pos=0, final_scores=0.4074089,
final_local_scores=0.33255213, pem=0.0, similarity=14.013732]
10: 7191778 Russian_Trading_System
    [pem_pos=0, final_scores=0.47973394,
final_local_scores=0.5896165, pem=0.0, similarity=21.685045]
11: 320489 Gazprom
    [pem_pos=0, final_scores=0.48421007,
final_local_scores=0.4164458, pem=0.0, similarity=16.517286]
```

Sample: b'844_Serbian'

[1 1 Serbian 1] policeman shot dead in [2 2 Kosovo 2]
province . [3 3 BELGRADE 3] 1996-08-29 A policeman has been
shot dead in [4 4 Serbia 4] 's troubled [5 5 Kosovo 5]
province , [6 6 Serbian 6] police said on Thursday . It was
the fifth attack on police this month in the southern
province , a hot spot of ethnic tension where the [7 7 Albani
an 7] majority have boycotted [8 8 Serbian 8] institutions
and set up their own , which are considered illegal by [9 9 Be
lgrade 9] . The slain policeman Ejup Bajgora , 42 , was an [10
10 Albanian 10] who had served in the [11 11 Serbian 11]
police and state security since 1987 , police told the [12 12
Yugoslav 12] news agency [13 13 Tanjug 13] . He was shot on
Wednesday afternoon as he stepped off a bus near his family
home in the village of Donje Ljupce in the municipality of [14
14 Podujevo 14] . Just hours before Wednesday 's shooting ,
three hand grenades were thrown at a police station in [15 15
Celopek 15] . They caused damage but no casualties , police
said . The [16 16 Serbian 16] authorities blame [17 17 Albani
an 17] dissidents for the recent spate of attacks . None of
the attackers has been caught . [18 18 Kosovo 18] 's autonomy
was revoked in 1987 and [19 19 Serb 19] police forces cracked
down on [20 20 Albanian 20] protests . [21 21 Albanian 21]
moderates want autonomy restored but hardliners want to join
up with neighbouring [22 22 Albania 22] . The [23 23 Serbs 23]
, who make up 10 percent of the province 's 1.8 million
people , claim [24 24 Kosovo 24] as the cradle of their
culture .

```
1: 29265 Serbia
2: 17391 Kosovo
3: 55904 Belgrade
4: 29265 Serbia
5: 17391 Kosovo
6: 29265 Serbia
7: 738 Albania
8: 29265 Serbia
9: 55904 Belgrade
10: 738 Albania
11: 29265 Serbia
12: 34244 Yugoslavia
13: 4151905 Tanjug
14: 777359 Podujevo
15: 14431526 Čelopek,_Brvenica
16: 29265 Serbia
17: 738 Albania
18: 17391 Kosovo
19: 29265 Serbia
20: 738 Albania
21: 738 Albania
22: 738 Albania
23: 29265 Serbia
24: 17391 Kosovo
1: 29265 Serbia
    [pem_pos=1, final_scores=0.58040076,
final_local_scores=0.52513826, pem=-1.2504034,
similarity=26.560774]
2: 17391 Kosovo
    [pem_pos=0, final_scores=0.38730735,
final_local_scores=0.32706916, pem=-0.3821891,
similarity=15.928513]
3: 55904 Belgrade
    [pem_pos=0, final_scores=0.4957283,
final_local_scores=0.45978463, pem=-0.05193398,
similarity=18.093029]
4: 29265 Serbia
    [pem_pos=0, final_scores=0.603243,
final_local_scores=0.5500899, pem=-0.34077588,
similarity=22.358686]
5: 17391 Kosovo
    [pem_pos=0, final_scores=0.4031291,
final_local_scores=0.344352, pem=-0.3821891,
similarity=16.444267]
6: 29265 Serbia
    [pem_pos=1, final_scores=0.59844035,
final_local_scores=0.54484373, pem=-1.2504034,
similarity=27.148825]
7: 738 Albania
    [pem_pos=2, final_scores=0.6178773,
final_local_scores=0.5738448, pem=-1.5560378,
similarity=29.676361]
```

8: 29265 Serbia
[pem_pos=1, final_scores=0.63886654,
final_local_scores=0.5890032, pem=-1.2504034,
similarity=28.466633]
9: 55904 Belgrade
[pem_pos=0, final_scores=0.4148925,
final_local_scores=0.3714838, pem=-0.05193398,
similarity=15.457955]
10: 738 Albania
[pem_pos=2, final_scores=0.5328658,
final_local_scores=0.4809826, pem=-1.5560378,
similarity=26.905169]
11: 29265 Serbia
[pem_pos=1, final_scores=0.62670016,
final_local_scores=0.5757133, pem=-1.2504034,
similarity=28.070034]
12: 34244 Yugoslavia
[pem_pos=0, final_scores=0.39741957,
final_local_scores=0.34734404, pem=-0.7596219,
similarity=18.586094]
13: 4151905 Tanjug
[pem_pos=0, final_scores=0.47789073,
final_local_scores=0.57442373, pem=0.0, similarity=21.231663]
14: 777359 Podujevo
[pem_pos=0, final_scores=0.478316,
final_local_scores=0.5189915, pem=0.0, similarity=19.577454]
15: 14431526 Čelopek,_Brvenica
[pem_pos=1, final_scores=0.20217131,
final_local_scores=0.41874397, pem=-2.7080503,
similarity=31.312668]
16: 29265 Serbia
[pem_pos=1, final_scores=0.7205551,
final_local_scores=0.6782356, pem=-1.2504034,
similarity=31.129503]
17: 738 Albania
[pem_pos=2, final_scores=0.71229345,
final_local_scores=0.6769801, pem=-1.5560378,
similarity=32.754124]
18: 17391 Kosovo
[pem_pos=0, final_scores=0.38872445,
final_local_scores=0.3286171, pem=-0.3821891,
similarity=15.974706]
19: 29265 Serbia
[pem_pos=1, final_scores=0.53666145,
final_local_scores=0.47735965, pem=-2.1726613,
similarity=30.150347]
20: 738 Albania
[pem_pos=2, final_scores=0.69079846,
final_local_scores=0.6535, pem=-1.5560378,
similarity=32.053432]
21: 738 Albania
[pem_pos=2, final_scores=0.6263401,

```
final_local_scores=0.58308905, pem=-1.5560378,
similarity=29.952229]
22: 738 Albania
    [pem_pos=0, final_scores=0.6077209,
final_local_scores=0.56275046, pem=-0.41162592,
similarity=23.121794]
23: 29265 Serbia
    [pem_pos=4, final_scores=0.43146348,
final_local_scores=0.36244687, pem=-2.3277044,
similarity=27.564268]
24: 17391 Kosovo
    [pem_pos=0, final_scores=0.36353666,
final_local_scores=0.3011033, pem=-0.3821891,
similarity=15.15364]

Sample: b'845_Viral'
Viral meningitis epidemic kills 10 in [1 1 Romania 1] . [2 2
BUCHAREST 2] 1996-08-29 Viral meningitis has killed 10 people
in [3 3 Romania 3] 's capital [4 4 Bucharest 4] this month in
what doctors said on Thursday was the worst epidemic of its
type in the country for a decade . Some 170 middle-aged and
elderly people with the disease were being treated in
hospital , doctors said . Doctor Emanuil Ceausu , head of [5 5
Bucharest 5] 's [6 6 Victor Babes 6] hospital for infectious
diseases , said the epidemic had been caused by a virus yet
to be identified . Illness from viral meningitis lasts around
a week . It affects the gastro-intestinal tract , causing
high fever , headache and vomiting . In 1986 [7 7 Romania 7]
suffered an epidemic of the more dangerous bacterial
meningitis which has killed some 15,000 people in central [8 8
Africa 8] this year .
1: 25445 Romania
2: 36877 Bucharest
3: 25445 Romania
4: 36877 Bucharest
5: 36877 Bucharest
6: 495612 Victor_Babes
7: 25445 Romania
8: 5334607 Africa
1: 25445 Romania
    [pem_pos=0, final_scores=0.74541956,
final_local_scores=0.711482, pem=-0.21241479,
similarity=26.476898]
2: 36877 Bucharest
    [pem_pos=0, final_scores=0.52796876,
final_local_scores=0.47828066, pem=-0.01781679,
similarity=18.459452]
3: 25445 Romania
    [pem_pos=0, final_scores=0.7391497,
final_local_scores=0.7046331, pem=-0.21241479,
similarity=26.272514]
4: 36877 Bucharest
```

```
[pem_pos=0, final_scores=0.57936054,
final_local_scores=0.5344184, pem=-0.01781679,
similarity=20.134714]
5: 36877 Bucharest
    [pem_pos=0, final_scores=0.64980966,
final_local_scores=0.61137336, pem=-0.01781679,
similarity=22.431204]
6: 495612 Victor Babeş
    [pem_pos=0, final_scores=0.49582094,
final_local_scores=0.6460374, pem=0.0, similarity=23.368757]
7: 25445 Romania
    [pem_pos=0, final_scores=0.7250625,
final_local_scores=0.689245, pem=-0.21241479,
similarity=25.813303]
8: 5334607 Africa
    [pem_pos=0, final_scores=0.5225096,
final_local_scores=0.59552586, pem=-0.13197383,
similarity=22.579086]
```

Sample: b'846 Serbia'

[1 1 Serbia 1] 's [2 2 Zastava 2] workers protest enters 9th day . [3 3 BELGRADE 3] 1996-08-29 Workers at [4 4 Serbia 4] 's [5 5 Zastava 5] arms factory entered the ninth day of their protest over unpaid wages on Thursday with management accused them of rejecting talks . " The workers keep on gathering in the centre of the town , " the factory 's general manager Vukasin Filipovic told [6 6 Reuters 6] . " But they do not want to talk to anyone . " " They want to discuss in public , at their protest meetings , " Filipovic said . " And that is impossible . " The [7 7 Zastava 7] works in the central town of [8 8 Kragujevac 8] is the backbone of [9 9 Serbia 9] 's defence industry , supplying the army with a whole range of weapons . Its workers are staging protests in the town 's main square demanding June and July wages and last year 's holiday pay . On Wednesday , the union demanded the resignation of the factory manager . But Filipovic said he would not quit under pressure . " We can talk about it and I am prepared to take all the consequences of mismanagement if any . "

```
1: 29265 Serbia
2: 8370927 Zastava_Arms
3: 55904 Belgrade
4: 29265 Serbia
5: 8370927 Zastava_Arms
6: 18998750 Reuters
7: 8370927 Zastava_Arms
8: 147111 Kragujevac
9: 29265 Serbia
1: 29265 Serbia
    [pem_pos=0, final_scores=0.5765272,
final_local_scores=0.5599692, pem=-0.34077588,
similarity=22.653503]
```

```
2: 8370927 Zastava_Arms
    [pem_pos=1, final_scores=0.5087561,
final_local_scores=0.552306, pem=-0.64282215,
similarity=24.06739]
3: 55904 Belgrade
    [pem_pos=0, final_scores=0.5119509,
final_local_scores=0.4990855, pem=-0.05193398,
similarity=19.265844]
4: 29265 Serbia
    [pem_pos=0, final_scores=0.6154682,
final_local_scores=0.6025064, pem=-0.34077588,
similarity=23.9229]
5: 8370927 Zastava_Arms
    [pem_pos=1, final_scores=0.49445802,
final_local_scores=0.53668755, pem=-0.64282215,
similarity=23.601305]
6: 18998750 Reuters
    [pem_pos=0, final_scores=0.30182898,
final_local_scores=0.38886112, pem=0.0, similarity=15.694105]
7: 8370927 Zastava_Arms
    [pem_pos=1, final_scores=0.53803533,
final_local_scores=0.58428913, pem=-0.64282215,
similarity=25.021832]
8: 147111 Kragujevac
    [pem_pos=0, final_scores=0.45076376,
final_local_scores=0.45099455, pem=0.0, similarity=17.54829]
9: 29265 Serbia
    [pem_pos=0, final_scores=0.6126322,
final_local_scores=0.5994084, pem=-0.34077588,
similarity=23.830448]
```

Sample: b'847_Venezuela'

```
[1 1 Venezuela 1] FinMin to make statement midday Thursday .
[2 2 CARACAS 2] 1996-08-28 [3 3 Venezuelan 3] Finance
Minister Luis Raul Matos Azocar will make an " important
announcement " on Thursday at 1230 local / 1630 [4 4 GMT 4]
at a Central Bank press conference , the Finance Ministry
said . The press conference replaces Matos 's scheduled
appearance at an IMF-hosted seminar Thursday on [5 5 Venezuela
5] 's economic reform program , Venezuelan Agenda . -- [6 6 Ca
racas 6] newsroom , 582 834405
1: 32374 Venezuela
2: 55289 Caracas
3: 32374 Venezuela
4: 12701 Greenwich_Mean_Time
5: 32374 Venezuela
6: 55289 Caracas
1: 32374 Venezuela
    [pem_pos=0, final_scores=0.55131036,
final_local_scores=0.5114303, pem=-0.21226841,
similarity=20.506163]
2: 55289 Caracas
```

```
[pem_pos=0, final_scores=0.49049884,
final_local_scores=0.45404404, pem=0.0, similarity=17.639294]
3: 32374 Venezuela
    [pem_pos=0, final_scores=0.7474474,
final_local_scores=0.72568023, pem=-0.11529429,
similarity=26.372448]
4: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.59981346,
final_local_scores=0.7464163, pem=0.0, similarity=26.364264]
5: 32374 Venezuela
    [pem_pos=0, final_scores=0.5648718,
final_local_scores=0.5262441, pem=-0.21226841,
similarity=20.948236]
6: 55289 Caracas
    [pem_pos=0, final_scores=0.58237976,
final_local_scores=0.5544099, pem=0.0, similarity=20.63441]
```

Sample: b'848_Colombia'

```
[1 1 Colombia 1] police find marijuana on ship . [2 2 BOGOTA
2] , [3 3 Colombia 3] 1996-08-29 Police said they found 35
metric tons of marijuana on Thursday on a ship preparing to
set sail for the [4 4 Netherlands 4] from [5 5 Colombia 5]'s
[6 6 Caribbean 6] port of [7 7 Cartagena 7] . They said the
drug had been packed into a shipping container and was
surrounded by ground coffee . No arrests had been made , a
police spokesman said .
```

```
1: 5222 Colombia
2: 211271 Bogotá
3: 5222 Colombia
4: 21148 Netherlands
5: 5222 Colombia
6: 18956035 Caribbean
7: 63927 Cartagena,_Colombia
1: 5222 Colombia
    [pem_pos=0, final_scores=0.7827535,
final_local_scores=0.75476575, pem=-0.22328302,
similarity=27.827677]
2: 211271 Bogotá
    [pem_pos=0, final_scores=0.6122519,
final_local_scores=0.60538155, pem=0.0, similarity=22.155506]
3: 5222 Colombia
    [pem_pos=0, final_scores=0.6134707,
final_local_scores=0.56985, pem=-0.22328302,
similarity=22.309422]
4: 21148 Netherlands
    [pem_pos=0, final_scores=0.44207704,
final_local_scores=0.46322644, pem=-0.24940795,
similarity=19.269634]
5: 5222 Colombia
    [pem_pos=0, final_scores=0.59768665,
final_local_scores=0.5526084, pem=-0.22328302,
similarity=21.794897]
```

```
6: 18956035 Caribbean
   [pem_pos=0, final_scores=0.36049658,
final_local_scores=0.3428708, pem=-0.09554641,
similarity=14.841258]
7: 63927 Cartagena,_Colombia
   [pem_pos=0, final_scores=0.2903428,
final_local_scores=0.2764993, pem=-0.41457057,
similarity=14.595503]

Sample: b'849_Venezuela'
[1 1 Venezuela 1] non-oil exports rise 10.6 pct in July . [2
2 CARACAS 2] 1996-08-29 [3 3 Venezuela 3] 's non-traditional
exports , which exclude oil and iron , rose 10.6 percent in
July to reach $ 334 million compared to $ 302 million in
June , according to the Central Office of Statistics and
Information ( OCEI ) . " The rise was due to the end of
exchange controls , " OCEI said . Foreign exchange controls
were removed April 22 as part of a wider IMF-sponsored
program . Nevertheless , exports over the first seven months
of the year were 16.8 percent lower than during the same
period last year , at $ 2.240 billion compared to $ 2.693
billion . Over the seven months , the private sector
accounted for 76 percent of total exports , with " common
metals " the strongest export sector accounting for $ 951
million , or 42.5 percent of total exports . " Chemical
products " came next with a 13 percent share , then "
transport materials " with nine percent , and finally foods ,
drinks and tobacco with 6.3 percent . [4 4 Colombia 4] was
the chief market for [5 5 Venezuela 5] 's non-traditional
exports with 27.4 percent . The [6 6 U.S. 6] followed with a
24.6 percent share . -- [7 7 Caracas 7] newsroom , 582 834405
REUTER JPR
1: 32374 Venezuela
2: 55289 Caracas
3: 32374 Venezuela
4: 5222 Colombia
5: 32374 Venezuela
6: 3434750 United_States
7: 55289 Caracas
1: 32374 Venezuela
   [pem_pos=0, final_scores=0.7871683,
final_local_scores=0.7488021, pem=-0.21226841,
similarity=27.589811]
2: 55289 Caracas
   [pem_pos=0, final_scores=0.4888224,
final_local_scores=0.45422155, pem=0.0, similarity=17.64459]
3: 32374 Venezuela
   [pem_pos=0, final_scores=0.69635755,
final_local_scores=0.64960504, pem=-0.21226841,
similarity=24.629572]
4: 5222 Colombia
   [pem_pos=0, final_scores=0.58446515,
```

```
final_local_scores=0.5410865, pem=-0.22328302,  
similarity=21.451061]  
5: 32374 Venezuela  
    [pem_pos=0, final_scores=0.62356,  
final_local_scores=0.5700848, pem=-0.21226841,  
similarity=22.25653]  
6: 3434750 United_States  
    [pem_pos=0, final_scores=0.464168,  
final_local_scores=0.47788495, pem=-0.14075686,  
similarity=19.116209]  
7: 55289 Caracas  
    [pem_pos=0, final_scores=0.56882936,  
final_local_scores=0.541617, pem=0.0, similarity=20.252644]
```

Sample: b'850_Buenos'

[1 1 Buenos Aires 1] fraud cops held in extortion racket . [2
2 BUENOS AIRES 2] , [3 3 Argentina 3] 1996-08-29 Thirteen
senior police officers from the fraud squad of [4 4 Buenos Ai
res 4] province have been arrested on charges of running an
extortion racket , security officials said on Thursday . They
included all the top officers from the fraud division of the
north of [5 5 Buenos Aires 5] province , including
Commissioner Juan Carlos Lago . Police were seeking a 14th
officer . [6 6 La Nacion 6] newspaper said the officers were
suspected of demanding bribes of \$ 50,000 to \$ 500,000 from
companies being investigated for tax evasion in order to "
lose " their files . The credibility of the [7 7 Buenos Aires
7] provincial police , the largest force in [8 8 Argentina 8]
, has been undermined this year by scandals that included the
indictment of three officers for links to the 1994 bombing of
a [9 9 Jewish 9] community centre and the arrest of an entire
drugs squad for drug trafficking . Alberto Piotti , security
chief of [10 10 Buenos Aires 10] province , told local
television that 3,600 dishonest officers had been purged from
the force 's ranks in the past five years . " It is an
ongoing task . And these investigations into police
corruption are only possible because there are people brave
enough to denounce them , " Piotti said , promising a major
overhaul of the provincial police next month .

```
1: 255919 Buenos_Aires  
2: 255919 Buenos_Aires  
3: 18951905 Argentina  
4: 255932 Buenos_Aires_Province  
5: 255932 Buenos_Aires_Province  
6: 2122632 La_Nación  
7: 255932 Buenos_Aires_Province  
8: 18951905 Argentina  
9: 25955086 Jews  
10: 255932 Buenos_Aires_Province  
1: 255919 Buenos_Aires  
    [pem_pos=0, final_scores=0.3206988,  
final_local_scores=0.25725263, pem=-0.12597719,
```

```
similarity=12.451728]
2: 255919 Buenos_Aires
    [pem_pos=0, final_scores=0.47462487,
final_local_scores=0.42539346, pem=-0.12597719,
similarity=17.469385]
3: 18951905 Argentina
    [pem_pos=0, final_scores=0.61397403,
final_local_scores=0.59181327, pem=-0.30003807,
similarity=23.382254]
4: 255932 Buenos_Aires_Province
    [pem_pos=1, final_scores=0.40297037,
final_local_scores=0.34216422, pem=-2.416869,
similarity=27.443886]
5: 255932 Buenos_Aires_Province
    [pem_pos=1, final_scores=0.3954044,
final_local_scores=0.3338995, pem=-2.416869,
similarity=27.197248]
6: 2122632 La_Nación
    [pem_pos=0, final_scores=0.43098485,
final_local_scores=0.49173707, pem=0.0, similarity=18.764128]
7: 255932 Buenos_Aires_Province
    [pem_pos=1, final_scores=0.38583195,
final_local_scores=0.32344306, pem=-2.416869,
similarity=26.885208]
8: 18951905 Argentina
    [pem_pos=0, final_scores=0.49350697,
final_local_scores=0.4602213, pem=-0.30003807,
similarity=19.455288]
9: 25955086 Jews
    [pem_pos=1, final_scores=0.38707983,
final_local_scores=0.46295893, pem=-0.80476934,
similarity=22.281792]
10: 255932 Buenos_Aires_Province
    [pem_pos=1, final_scores=0.39080614,
final_local_scores=0.3288766, pem=-2.416869,
similarity=27.047356]
```

Sample: b'851_Brazil'
[1 1 Brazil 1] gov't set to send 1997 budget to [2 2 Congress
2] . [3 3 BRASILIA 3] 1996-08-29 [4 4 Brazilian 4] Planning
Minister Antonio Kandir will submit to a draft copy of the
1997 federal budget to [5 5 Congress 5] on Thursday , a
ministry spokeswoman said . [6 6 Congress 6] is
constitutionally obliged to approve the budget by the end of
year but regularly fails to meet that requirement .

```
1: 3383 Brazil
2: 690287 National_Congress_of_Brazil
3: 4752 Brasília
4: 3383 Brazil
5: 690287 National_Congress_of_Brazil
6: 690287 National_Congress_of_Brazil
1: 3383 Brazil
```

```
[pem_pos=0, final_scores=0.6463614,
final_local_scores=0.6879681, pem=-0.21313342,
similarity=25.779104]
2: 690287 National_Congress_of_Brazil
    [pem_pos=7, final_scores=0.5265569,
final_local_scores=0.514866, pem=-3.7206802,
similarity=39.68798]
3: 4752 Brasília
    [pem_pos=0, final_scores=0.4497692,
final_local_scores=0.4486912, pem=0.0, similarity=17.479553]
4: 3383 Brazil
    [pem_pos=0, final_scores=0.65909284,
final_local_scores=0.70187515, pem=-0.2733083,
similarity=26.52136]
5: 690287 National_Congress_of_Brazil
    [pem_pos=7, final_scores=0.53520507,
final_local_scores=0.52431273, pem=-3.7206802,
similarity=39.96989]
6: 690287 National_Congress_of_Brazil
    [pem_pos=7, final_scores=0.40115285,
final_local_scores=0.37788105, pem=-3.7206802,
similarity=35.600075]
```

Sample: b'852_Mexico'

[1 1 Mexico 1] same-day Cetes rates rise on nervousness . [2 2 MEXICO CITY 2] 1996-08-29 [3 3 Mexico 3] 's same-day Cetes rates rose 50 basis points to 24.25 percent on nervousness over a new round of attacks by guerrillas in two southern states , dealers said . " There are people who are taking advantage of the news to put pressure on rates , however , there are enough players who will buy and that will keep rates from rising too much , " said one dealer . Co-ordinated guerrilla attacks in two southern states overnight that left at least 13 people dead have caused nervousness in [4 4 Mexico 4] 's financial markets . Bank notes and acceptances , including pagares , rose 46 basis points to 25.10 percent . Dealers said that the volume of longer-term government paper declined due to market nervousness . At least 13 people were killed when scores of masked rebels struck at police and military posts in [5 5 Oaxaca 5] and [6 6 Guerrero 6] states overnight in the biggest assaults in more than two years , officials said on Thursday . Maturing credits are seen at 2.209 billion pesos , and there is an oversupply of 684 billion pesos from the primary auction . Dealers estimate that the shortfall will increase due to the inflow of funds before the end of the month . --- Patricia Lezama , [7 7 Mexico City 7] newroom (525) 728 9554

```
1: 3966054 Mexico
2: 18987 Mexico_City
3: 3966054 Mexico
4: 3966054 Mexico
5: 156896 Oaxaca
```

```
6: 195816 Guerrero
7: 18987 Mexico_City
1: 3966054 Mexico
    [pem_pos=0, final_scores=0.76911217,
final_local_scores=0.72855437, pem=-0.18258883,
similarity=26.824175]
2: 18987 Mexico_City
    [pem_pos=0, final_scores=0.4244663,
final_local_scores=0.39018643, pem=-0.028392298,
similarity=15.888057]
3: 3966054 Mexico
    [pem_pos=0, final_scores=0.8091535,
final_local_scores=0.7722934, pem=-0.18258883,
similarity=28.129435]
4: 3966054 Mexico
    [pem_pos=0, final_scores=0.6214399,
final_local_scores=0.56316227, pem=-0.18258883,
similarity=21.888544]
5: 156896 Oaxaca
    [pem_pos=0, final_scores=0.3395797,
final_local_scores=0.3184639, pem=0.0, similarity=13.59331]
6: 195816 Guerrero
    [pem_pos=0, final_scores=0.57704055,
final_local_scores=0.56988794, pem=-0.22545685,
similarity=22.322374]
7: 18987 Mexico_City
    [pem_pos=0, final_scores=0.43734032,
final_local_scores=0.40424937, pem=-0.028392298,
similarity=16.307722]
```

Sample: b'853_Tension'

Tension builds in [1 1 Mexican 1] state ahead of elections .
[CORRECTED 05:53 [2 2 GMT 2]] [3 3 CHILPANCINGO 3] , [4 4 Mexico 4] 1996-08-28 Pre-electoral bickering flared on Wednesday in the troubled western [5 5 Mexican 5] state of [6 6 Guerrero 6] as some opposition politicians demanded the army pull out of the area ahead of an upcoming state poll . The mayor of [7 7 Acatepec 7] , a small town some 310 miles (500 km) south of [8 8 Mexico City 8] , sent a letter to [9 9 Mexico 9] 's [10 10 National Human Rights Commission 10] complaining the army 's heavy presence in the town would interfere with the Oct. 6 election . Mayor Antonio Gonzalez Garcia , of the opposition Revolutionary Workers ' Party , said in Wednesday 's letter that army troops recently raided several local farms , stole cattle and raped women . The letter was signed by some 200 area residents and indigenous leaders . Some electoral watchdog groups also said the presence of the army , which has fanned out across the state in the past month in search of a new guerrilla group , was likely to intimidate voters and had stirred up tension in the state . A [11 11 French 11] group of electoral observers , Agir Ensemble pour les Droits de l'Homme , concluded the army

presence exerted a heavy psychological pressure on local farmers and would prevent a fair vote . Up for grabs in the election are the state legislature and 75 town halls . (Corrects to show elections are not for governor) . Despite the criticism , acting state Gov . Angel Aguirre pledged the elections will be free and fair and said he did not expect any trouble from the elusive new guerrilla group , the [12 12 Popular Revolutionary Army 12] . " The electoral process has been proceeding in accordance with the new state electoral law , " Aguirre said , adding that the poll would be " an exercise in true democracy . "

1: 3966054 Mexico
2: 12701 Greenwich_Mean_Time
3: 195822 Chilpancingo
4: 3966054 Mexico
5: 3966054 Mexico
6: 195816 Guerrero
7: 20822084 Acatepec
8: 18987 Mexico_City
9: 3966054 Mexico
10: 2685480 National_Human_Rights_Commission_(Mexico)
11: 5843419 France
12: 1529105 Popular_Revolutionary_Army
1: 3966054 Mexico
[pem_pos=0, final_scores=0.5738687,
final_local_scores=0.5360203, pem=-0.3968012,
similarity=22.243494]
2: 12701 Greenwich_Mean_Time
[pem_pos=0, final_scores=0.5512568,
final_local_scores=0.67943454, pem=0.0, similarity=24.365395]
3: 195822 Chilpancingo
[pem_pos=0, final_scores=0.45395404,
final_local_scores=0.50419337, pem=0.0, similarity=19.135849]
4: 3966054 Mexico
[pem_pos=0, final_scores=0.7250219,
final_local_scores=0.7011321, pem=-0.18258883,
similarity=26.00584]
5: 3966054 Mexico
[pem_pos=0, final_scores=0.4749021,
final_local_scores=0.42791426, pem=-0.3968012,
similarity=19.017395]
6: 195816 Guerrero
[pem_pos=0, final_scores=0.51293916,
final_local_scores=0.5178977, pem=-0.22545685,
similarity=20.770884]
7: 20822084 Acatepec
[pem_pos=0, final_scores=0.1721025,
final_local_scores=0.5225549, pem=0.0, similarity=19.683792]
8: 18987 Mexico_City
[pem_pos=0, final_scores=0.47322983,
final_local_scores=0.44944245, pem=-0.028392298,
similarity=17.656374]

```
9: 3966054 Mexico
    [pem_pos=0, final_scores=0.7084066,
final_local_scores=0.6829825, pem=-0.18258883,
similarity=25.46422]
10: 2685480 National_Human_Rights_Commission_(Mexico)
    [pem_pos=0, final_scores=0.43159354,
final_local_scores=0.5313075, pem=-0.21949057,
similarity=21.138613]
11: 5843419 France
    [pem_pos=0, final_scores=0.53702945,
final_local_scores=0.5826553, pem=-0.80223274,
similarity=25.839975]
12: 1529105 Popular_Revolutionary_Army
    [pem_pos=0, final_scores=0.47937822,
final_local_scores=0.6205246, pem=0.0, similarity=22.607405]
```

Sample: b'854_Brazil'

[1 1 Brazil 1] police arrest wanted [2 2 Italian 2] man - report . [3 3 SAO PAULO 3] , [4 4 Brazil 4] 1996-08-28 [5 5 Brazilian 5] authorities on Wednesday arrested a 47-year-old [6 6 Italian 6] man wanted in [7 7 Italy 7] for ties to the leftist [8 8 Red Brigade 8] guerrilla group of the 1970s , local television said . [9 9 TV Globo 9] said the [10 10 Supreme Federal Tribunal 10] ordered the arrest of Luciano Pessina , a political scientist who owns two [11 11 Rio de Janeiro 11] restaurants , based on an extradition request from the [12 12 Italian 12] government . The report , which could not be independently verified on Wednesday night , said Pessina was sentenced in [13 13 Italy 13] to eight years and 11 months in prison for robbery and illegal weapons and explosives possession . [14 14 Globo 14] quoted Pessina 's lawyer as saying he had already been imprisoned in [15 15 Italy 15] and , when freed , travelled to [16 16 Brazil 16] .

```
1: 3383 Brazil
2: 14532 Italy
3: 390875 São_Paulo
4: 3383 Brazil
5: 3383 Brazil
6: 14532 Italy
7: 14532 Italy
8: 41249202 Red_Brigades
9: 772540 Rede_Globo
10: 3359076 Supreme_Federal_Court
11: 25936 Rio_de_Janeiro
12: 14532 Italy
13: 14532 Italy
14: 772540 Rede_Globo
15: 14532 Italy
16: 3383 Brazil
1: 3383 Brazil
    [pem_pos=0, final_scores=0.77983814,
final_local_scores=0.7721174, pem=-0.21313342,
```

```
similarity=28.290287]
2: 14532 Italy
    [pem_pos=0, final_scores=0.76289946,
final_local_scores=0.7439025, pem=-0.75948656,
similarity=30.419455]
3: 390875 São_Paulo
    [pem_pos=0, final_scores=0.41194695,
final_local_scores=0.3795231, pem=0.0, similarity=15.41544]
4: 3383 Brazil
    [pem_pos=0, final_scores=0.8018841,
final_local_scores=0.79619926, pem=-0.21313342,
similarity=29.00894]
5: 3383 Brazil
    [pem_pos=0, final_scores=0.754309,
final_local_scores=0.74423075, pem=-0.2733083,
similarity=27.785336]
6: 14532 Italy
    [pem_pos=0, final_scores=0.71614236,
final_local_scores=0.69282746, pem=-0.75948656,
similarity=28.895271]
7: 14532 Italy
    [pem_pos=0, final_scores=0.741126,
final_local_scores=0.7201183, pem=-0.20612453,
similarity=26.700417]
8: 41249202 Red_Brigades
    [pem_pos=0, final_scores=0.48181337,
final_local_scores=0.58199733, pem=0.0, similarity=21.457674]
9: 772540 Rede_Globo
    [pem_pos=0, final_scores=0.4694202,
final_local_scores=0.50904256, pem=0.0, similarity=19.280558]
10: 3359076 Supreme_Federal_Court
    [pem_pos=0, final_scores=0.2975363,
final_local_scores=0.39824438, pem=0.0, similarity=15.97412]
11: 25936 Rio_de_Janeiro
    [pem_pos=0, final_scores=0.37372297,
final_local_scores=0.33010772, pem=-0.1260924,
similarity=14.626496]
12: 14532 Italy
    [pem_pos=0, final_scores=0.7912802,
final_local_scores=0.7749042, pem=-0.75948656,
similarity=31.344606]
13: 14532 Italy
    [pem_pos=0, final_scores=0.7596032,
final_local_scores=0.7403018, pem=-0.20612453,
similarity=27.302732]
14: 772540 Rede_Globo
    [pem_pos=0, final_scores=0.34838933,
final_local_scores=0.37683475, pem=0.0, similarity=15.335214]
15: 14532 Italy
    [pem_pos=0, final_scores=0.7257685,
final_local_scores=0.7033426, pem=-0.20612453,
similarity=26.199797]
```

```
16: 3383 Brazil
    [pem_pos=0, final_scores=0.70404804,
final_local_scores=0.6893282, pem=-0.21313342,
similarity=25.819693]
```

Sample: b'855_Seven'

Seven churches slam [1 1 Brazil 1] rural violence , impunity . [2 2 BRASILIA 2] 1996-08-28 Seven churches joined voices on Wednesday to condemn the " day-to-day violence " of [3 3 Brazil 3] 's rural hinterland and the government 's failure to punish those responsible for massacres of landless peasants . " In the name of [4 4 Jesus 4] , we want to bring your attention to what is going on in the [5 5 Brazilian 5] countryside , " two church umbrella groups said in a Letter to the [6 6 Brazilian 6] People . The letter from National Council of Christian Churches of Brazil and the Coordinate of Ecumenical Service was sent to President [7 7 Fernando Henrique Cardoso 7] after a seminar on endemic violence gripping rural [8 8 Brazil 8] . Thirty-six people have died so far this year in conflicts over land in the [9 9 Brazilian 9] countryside , including 19 landless peasants massacred by police in April in the northern state of [10 10 Para 10] . " The problem of land is one of the most serious facing [11 11 Brazil 11] , " said [12 12 Lucas Moreira Neves 12] , president of the [13 13 Catholic 13] church 's National Conference of Bishops of Brazil . The letter made reference to massacres of landless peasants in August 1995 and April 1996 , which claimed the lives of 27 landless peasants .

```
1: 3383 Brazil
```

```
2: 4752 Brasília
```

```
3: 3383 Brazil
```

```
4: 1095706 Jesus
```

```
5: 3383 Brazil
```

```
6: 3383 Brazil
```

```
7: 82491 Fernando_Henrique_Cardoso
```

```
8: 3383 Brazil
```

```
9: 3383 Brazil
```

```
10: 229521 Pará
```

```
11: 3383 Brazil
```

```
12: 8284063 Lucas_Moreira_Neves
```

```
13: 606848 Catholic_Church
```

```
1: 3383 Brazil
```

```
    [pem_pos=0, final_scores=0.71913165,
final_local_scores=0.67617697, pem=-0.21313342,
similarity=25.427235]
```

```
2: 4752 Brasília
```

```
    [pem_pos=0, final_scores=0.47208953,
final_local_scores=0.46665305, pem=0.0, similarity=18.015572]
```

```
3: 3383 Brazil
```

```
    [pem_pos=0, final_scores=0.7031155,
final_local_scores=0.6586818, pem=-0.21313342,
similarity=24.905144]
```

```
4: 1095706 Jesus
    [pem_pos=0, final_scores=0.47337717,
final_local_scores=0.5943692, pem=-0.0733903,
similarity=22.225983]
5: 3383 Brazil
    [pem_pos=0, final_scores=0.6536094,
final_local_scores=0.6046038, pem=-0.2733083,
similarity=23.61859]
6: 3383 Brazil
    [pem_pos=0, final_scores=0.5423848,
final_local_scores=0.48310786, pem=-0.2733083,
similarity=19.992908]
7: 82491 Fernando_Henrique_Cardoso
    [pem_pos=0, final_scores=0.50016934,
final_local_scores=0.56919706, pem=0.0, similarity=21.07569]
8: 3383 Brazil
    [pem_pos=0, final_scores=0.6651881,
final_local_scores=0.6172518, pem=-0.21313342,
similarity=23.66879]
9: 3383 Brazil
    [pem_pos=0, final_scores=0.6472897,
final_local_scores=0.59770054, pem=-0.2733083,
similarity=23.412582]
10: 229521 Pará
    [pem_pos=4, final_scores=0.39480436,
final_local_scores=0.39198107, pem=-1.817512,
similarity=25.671124]
11: 3383 Brazil
    [pem_pos=0, final_scores=0.5784669,
final_local_scores=0.5225221, pem=-0.21313342,
similarity=20.841867]
12: 8284063 Lucas_Moreira_Neves
    [pem_pos=0, final_scores=0.3879009,
final_local_scores=0.53474253, pem=0.0, similarity=20.047497]
13: 606848 Catholic_Church
    [pem_pos=1, final_scores=0.14480747,
final_local_scores=0.18271527, pem=-0.9041281,
similarity=14.45909]
```

Sample: b'856_Thai'

```
[1 1 Thai 1] official flees [2 2 Hong Kong 2] after passport
scam . [3 3 HONG KONG 3] 1996-08-29 A [4 4 Thai 4] consular
official has fled [5 5 Hong Kong 5] after being questioned by
anti-corruption police in connection with soliciting bribes
to issue a passport , [6 6 Hong Kong 6] government radio said
on Thursday . The unnamed suspect left the [7 7 British 7]
colony after being detained and then freed by the [8 8 Indepe
ndent Commission Against Corruption 8] ( [9 9 ICAC 9] ) , the
radio said . The man was released after his arrest on
Tuesday , pending further inquiries , the [10 10 ICAC 10]
said in a statement . The anti-graft body was discussing the
case with the [11 11 Thai 11] government , especially the
```

suspect 's status , it said . It is not clear if the fugitive had diplomatic status in [12 12 Hong_Kong 12] , and officials from the [13 13 Thai 13] Consulate were not available for comment . The arrest came after the [14 14 ICAC 14] received a complaint that the man had demanded a bribe of [15 15 HK\$ 15] 100,000 ([16 16 US\$ 16] 12,940) to issue a [17 17 Thai 17] passport , the [18 18 ICAC 18] said . At the time of his arrest , [19 19 ICAC 19] officers seized [20 20 HK\$ 20] 100,000 , it added . The [21 21 ICAC 21] has kept a close eye on passport scams after a [22 22 U.S. 22] official was jailed for trafficking fake [23 23 Honduran 23] passports as part of an immigration racket aimed at [24 24 Chinese 24] .

1: 30128 Thailand

2: 13404 Hong_Kong

3: 13404 Hong_Kong

4: 30128 Thailand

5: 13404 Hong_Kong

6: 13404 Hong_Kong

7: 31717 United_Kingdom

8: 534654

Independent_Commission_Against_Corruption_(Hong_Kong)

9: 534654

Independent_Commission_Against_Corruption_(Hong_Kong)

10: 534654

Independent_Commission_Against_Corruption_(Hong_Kong)

11: 30128 Thailand

12: 13404 Hong_Kong

13: 30128 Thailand

14: 534654

Independent_Commission_Against_Corruption_(Hong_Kong)

15: 309734 Hong_Kong_dollar

16: 18717338 United_States_dollar

17: 30128 Thailand

18: 534654

Independent_Commission_Against_Corruption_(Hong_Kong)

19: 534654

Independent_Commission_Against_Corruption_(Hong_Kong)

20: 309734 Hong_Kong_dollar

21: 534654

Independent_Commission_Against_Corruption_(Hong_Kong)

22: 3434750 United_States

23: 13394 Honduras

24: 372983 Chinese_people

1: 30128 Thailand

[pem_pos=2, final_scores=0.6945636,
final_local_scores=0.6916612, pem=-1.7504287,
similarity=34.249363]

2: 13404 Hong_Kong

[pem_pos=0, final_scores=0.5451501,
final_local_scores=0.49687243, pem=-0.22402541,
similarity=20.135662]

3: 13404 Hong_Kong

```
[pem_pos=0, final_scores=0.5716422,
final_local_scores=0.5258111, pem=-0.22402541,
similarity=20.99925]
4: 30128 Thailand
    [pem_pos=2, final_scores=0.6453221,
final_local_scores=0.6378723, pem=-1.7504287,
similarity=32.644196]
5: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5242589,
final_local_scores=0.474052, pem=-0.22402541,
similarity=19.454657]
6: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.51656634,
final_local_scores=0.465649, pem=-0.22402541,
similarity=19.203894]
7: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5095822,
final_local_scores=0.5207869, pem=-0.48936236,
similarity=22.292261]
8: 534654
Independent_Commission_Against_Corruption_(Hong_Kong)
    [pem_pos=0, final_scores=0.40194207,
final_local_scores=0.3954438, pem=0.0, similarity=15.890545]
9: 534654
Independent_Commission_Against_Corruption_(Hong_Kong)
    [pem_pos=0, final_scores=0.47039652,
final_local_scores=0.4702199, pem=0.0, similarity=18.122013]
10: 534654
Independent_Commission_Against_Corruption_(Hong_Kong)
    [pem_pos=0, final_scores=0.5071699,
final_local_scores=0.51038927, pem=0.0, similarity=19.320747]
11: 30128 Thailand
    [pem_pos=2, final_scores=0.736787,
final_local_scores=0.73778385, pem=-1.7504287,
similarity=35.625755]
12: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5321618,
final_local_scores=0.48268467, pem=-0.22402541,
similarity=19.712273]
13: 30128 Thailand
    [pem_pos=2, final_scores=0.72635245,
final_local_scores=0.72638565, pem=-1.7504287,
similarity=35.28561]
14: 534654
Independent_Commission_Against_Corruption_(Hong_Kong)
    [pem_pos=0, final_scores=0.51068187,
final_local_scores=0.51422554, pem=0.0, similarity=19.435228]
15: 309734 Hong_Kong_dollar
    [pem_pos=0, final_scores=0.6363058,
final_local_scores=0.65998363, pem=0.0, similarity=23.78494]
16: 18717338 United_States_dollar
    [pem_pos=0, final_scores=0.4861145,
```

```
final_local_scores=0.5186289, pem=0.0, similarity=19.566633]
17: 30128 Thailand
    [pem_pos=2, final_scores=0.6771623,
final_local_scores=0.6726529, pem=-1.7504287,
similarity=33.682117]
18: 534654
Independent_Commission_Against_Corruption_(Hong_Kong)
    [pem_pos=0, final_scores=0.4680603,
final_local_scores=0.467668, pem=0.0, similarity=18.045858]
19: 534654
Independent_Commission_Against_Corruption_(Hong_Kong)
    [pem_pos=0, final_scores=0.54412633,
final_local_scores=0.55075854, pem=0.0, similarity=20.525446]
20: 309734 Hong_Kong_dollar
    [pem_pos=0, final_scores=0.66978955,
final_local_scores=0.69655955, pem=0.0, similarity=24.87644]
21: 534654
Independent_Commission_Against_Corruption_(Hong_Kong)
    [pem_pos=0, final_scores=0.48371565,
final_local_scores=0.48476905, pem=0.0, similarity=18.556189]
22: 3434750 United_States
    [pem_pos=0, final_scores=0.50164396,
final_local_scores=0.5214953, pem=-0.14075686,
similarity=20.417631]
23: 13394 Honduras
    [pem_pos=0, final_scores=0.2801419,
final_local_scores=0.30965263, pem=-0.14737272,
similarity=14.131801]
24: 372983 Chinese_people
    [pem_pos=2, final_scores=0.16651769,
final_local_scores=0.16615331, pem=-2.5410082,
similarity=22.866457]
```

Sample: b'857_Thai'

```
[1 1 Thai 1] poll shows military wants PM Banharn out . [2 2
BANGKOK 2] 1996-08-29 [3 3 Thailand 3] 's powerful military
thinks the government is dishonest and Prime Minister [4 4 Ba
nharn Silpa-archa 4] 's resignation might solve the nation 's
political and economic woes , an opinion poll showed on
Thursday . Nearly half the 1,617 military personnel surveyed
in the Rajapat Institute poll suggested that Banharn resign ,
while about 28 percent thought he should dissolve parliament
and 24 percent thought a cabinet reshuffle could resolve the
government 's problems . Banharn , who leads a six-party , 13-
month-old coalition government , faces a no-confidence debate
in parliament next month . The prime minister , who has
already lost one coalition partner this month , is expected
to have a tough battle in the debate because of infighting in
his own party and warnings of more pullouts by other
coalition partners . The poll , conducted earlier this month
after Banharn 's coalition completed one year in office ,
showed the military would prefer General [5 5 Chatichai Choon
```

[havan](#) 5] -- a former prime minister who was ousted in a military coup in February 1991 -- as prime minister . Defence Minister [6 [6 Chavalit Yongchaiyudh](#) 6] , head of coalition member the [7 [7 New Aspiration Party](#) 7] , was the second choice for prime minister , the poll showed . Banharn came in last on the list of proposed leaders with less than one percent of the votes . Nearly two-thirds of the people surveyed thought the government was dishonest and insincere , and 65 percent blamed the government 's poor performance for the country 's economic slowdown . The opinion of the military in [8 [8 Thailand](#) 8] , which has seen 17 coups or attempted coups since the country switched to parliamentary democracy from absolute monarchy in 1932 , always carries weight in the political scene , despite officials ' vows to distance the military from politics .

```
1: 30128 Thailand
2: 56615 Bangkok
3: 30128 Thailand
4: 7147085 Banharn_Silpa-archa
5: 1863118 Chatichai_Choonhavan
6: 2014560 Chavalit_Yongchaiyudh
7: 2095269 New_Aspiration_Party
8: 30128 Thailand
1: 30128 Thailand
    [pem_pos=2, final_scores=0.6570137,
final_local_scores=0.6574622, pem=-1.7504287,
similarity=33.228798]
2: 56615 Bangkok
    [pem_pos=0, final_scores=0.47061223,
final_local_scores=0.47347426, pem=-0.0005963434,
similarity=18.222372]
3: 30128 Thailand
    [pem_pos=0, final_scores=0.7530815,
final_local_scores=0.76240164, pem=-0.24029273,
similarity=28.148046]
4: 7147085 Banharn_Silpa-archa
    [pem_pos=0, final_scores=0.60867125,
final_local_scores=0.67311513, pem=0.0, similarity=24.176811]
5: 1863118 Chatichai_Choonhavan
    [pem_pos=0, final_scores=0.74618894,
final_local_scores=0.82504123, pem=0.0, similarity=28.710588]
6: 2014560 Chavalit_Yongchaiyudh
    [pem_pos=0, final_scores=0.5730815,
final_local_scores=0.63072425, pem=0.0, similarity=22.911783]
7: 2095269 New_Aspiration_Party
    [pem_pos=0, final_scores=0.5926243,
final_local_scores=0.6870561, pem=0.0, similarity=24.592838]
8: 30128 Thailand
    [pem_pos=0, final_scores=0.5768778,
final_local_scores=0.5699258, pem=-0.24029273,
similarity=22.404184]
```

Sample: b'858_Hamas'
[1 1 Hamas 1] cleric jailed in [2 2 Israel 2] hospitalised briefly . [3 3 JERUSALEM 3] 1996-08-29 [4 4 Israeli 4] prison officials took jailed [5 5 Islamic 5] militant [6 6 Hamas 6] founder Sheikh [7 7 Ahmed Yassin 7] to hospital briefly on Thursday for medical tests , officials said . " This evening Sheikh [8 8 Yassin 8] completed medical checks and returned to [9 9 Ramle 9] prisons authority medical centre , " said a spokesman for [10 10 Israel 10] 's internal security ministry . A prison official said [11 11 Yassin 11] had a mild case of pneumonia . The 60-year-old [12 12 Moslem 12] cleric , jailed by [13 13 Israel 13] since 1989 , is serving a life sentence for ordering attacks by [14 14 Hamas 14] guerrillas against [15 15 Israeli 15] targets . The ailing [16 16 Yassin 16] is the spiritual leader of the fundamentalist [1 7 17 Hamas 17] group which has killed scores of [18 18 Israel is 18] in suicide attacks aimed at wrecking Israel-PL0 peace deals . [19 19 Palestinian 19] President [20 20 Yasser Arafat 20] has demanded that [21 21 Israel 21] release [22 22 Yassin 22] -- who is confined to a wheelchair -- on humanitarian grounds . [23 23 Israel 23] said last month after it recovered the body of a soldier abducted by [24 24 Hamas 24] seven years ago that it would consider freeing [25 25 Yassin 2 5] .

1: 13913 Hamas
2: 9282173 Israel
3: 16043 Jerusalem
4: 9282173 Israel
5: 6037917 Islam
6: 13913 Hamas
7: 80830 Ahmed_Yassin
8: 80830 Ahmed_Yassin
9: 323256 Ramla
10: 9282173 Israel
11: 80830 Ahmed_Yassin
12: 6037917 Islam
13: 9282173 Israel
14: 13913 Hamas
15: 9282173 Israel
16: 80830 Ahmed_Yassin
17: 13913 Hamas
18: 9282173 Israel
19: 241405 State_of_Palestine
20: 34289 Yasser_Arafat
21: 9282173 Israel
22: 80830 Ahmed_Yassin
23: 9282173 Israel
24: 13913 Hamas
25: 80830 Ahmed_Yassin
1: 13913 Hamas
[pem_pos=0, final_scores=0.44908482, final_local_scores=0.3982631, pem=0.0, similarity=15.974678]

2: 9282173 Israel
[pem_pos=1, final_scores=0.58950293,
final_local_scores=0.5690847, pem=-0.21279564,
similarity=22.229551]
3: 16043 Jerusalem
[pem_pos=0, final_scores=0.4965607,
final_local_scores=0.49627447, pem=-0.13239029,
similarity=19.619492]
4: 9282173 Israel
[pem_pos=0, final_scores=0.5298972,
final_local_scores=0.5039745, pem=-0.34790626,
similarity=21.021286]
5: 6037917 Islam
[pem_pos=0, final_scores=0.43221164,
final_local_scores=0.47399926, pem=-0.21141872,
similarity=19.384525]
6: 13913 Hamas
[pem_pos=0, final_scores=0.4467119,
final_local_scores=0.395671, pem=0.0, similarity=15.897325]
7: 80830 Ahmed_Yassin
[pem_pos=0, final_scores=0.6199334,
final_local_scores=0.5897919, pem=0.0, similarity=21.690279]
8: 80830 Ahmed_Yassin
[pem_pos=0, final_scores=0.6548032,
final_local_scores=0.6311178, pem=0.0, similarity=22.923527]
9: 323256 Ramla
[pem_pos=0, final_scores=0.42650908,
final_local_scores=0.45807493, pem=0.0, similarity=17.759583]
10: 9282173 Israel
[pem_pos=0, final_scores=0.6299565,
final_local_scores=0.61327404, pem=-0.21279564,
similarity=23.54825]
11: 80830 Ahmed_Yassin
[pem_pos=0, final_scores=0.7320452,
final_local_scores=0.7154929, pem=0.0, similarity=25.441448]
12: 6037917 Islam
[pem_pos=1, final_scores=0.43206376,
final_local_scores=0.47383773, pem=-1.2150909,
similarity=24.837828]
13: 9282173 Israel
[pem_pos=1, final_scores=0.6259067,
final_local_scores=0.6088503, pem=-0.21279564,
similarity=23.416235]
14: 13913 Hamas
[pem_pos=0, final_scores=0.45152837,
final_local_scores=0.4009323, pem=0.0, similarity=16.054333]
15: 9282173 Israel
[pem_pos=0, final_scores=0.5338682,
final_local_scores=0.50831217, pem=-0.34790626,
similarity=21.15073]
16: 80830 Ahmed_Yassin
[pem_pos=0, final_scores=0.6692399,

```
final_local_scores=0.64688766, pem=0.0, similarity=23.39413]
17: 13913 Hamas
    [pem_pos=0, final_scores=0.46068907,
final_local_scores=0.41093892, pem=0.0, similarity=16.35295]
18: 9282173 Israel
    [pem_pos=1, final_scores=0.40287632,
final_local_scores=0.36522347, pem=-0.5460739,
similarity=17.958344]
19: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.6369415,
final_local_scores=0.66495466, pem=-2.8508017,
similarity=39.43639]
20: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.60608774,
final_local_scores=0.5935288, pem=0.0, similarity=21.801796]
21: 9282173 Israel
    [pem_pos=1, final_scores=0.5872825,
final_local_scores=0.56665915, pem=-0.21279564,
similarity=22.157167]
22: 80830 Ahmed_Yassin
    [pem_pos=0, final_scores=0.74538416,
final_local_scores=0.73006374, pem=0.0, similarity=25.87627]
23: 9282173 Israel
    [pem_pos=1, final_scores=0.49187744,
final_local_scores=0.46244365, pem=-0.21279564,
similarity=19.04717]
24: 13913 Hamas
    [pem_pos=0, final_scores=0.48389602,
final_local_scores=0.436289, pem=0.0, similarity=17.109447]
25: 80830 Ahmed_Yassin
    [pem_pos=0, final_scores=0.5986929,
final_local_scores=0.5698258, pem=0.0, similarity=21.094452]
```

Sample: b'859_Moroccan'

[1 1 Moroccan 1] King meets former [2 2 Israel 2] PM [3 3 Pe
res 3] . [4 4 SKHIRAT 4] , [5 5 Morocco 5] 1996-08-29 [6 6 Ki
ng Hassan 6] of [7 7 Morocco 7] met former [8 8 Israeli 8]
prime minister [9 9 Shimon Peres 9] on Thursday at the
coastal resort of [10 10 Skhirat 10] , 20 km (12 miles)
south of [11 11 Rabat 11] , an official source said . " Mr [12
12 Shimon Peres 12] , who is on a purely private and family
visit to [13 13 Morocco 13] , was received on Thursday by his
Majesty [14 14 King Hassan 14] at the royal palace of [15 15 S
khirat 15] , " the source said . Accompanied by his wife and
grandson , [16 16 Peres 16] arrived in [17 17 Morocco 17] on
August 25 . [18 18 Peres 18] is expected to fly home on
Friday , officials said .

1: 19291 Morocco
2: 9282173 Israel
3: 48569 Shimon_Peres
4: 4512052 Skhirat
5: 19291 Morocco

```
6: 410355 Hassan_II_of_Morocco
7: 19291 Morocco
8: 9282173 Israel
9: 48569 Shimon_Peres
10: 4512052 Skhirat
11: 57400 Rabat
12: 48569 Shimon_Peres
13: 19291 Morocco
14: 410355 Hassan_II_of_Morocco
15: 4512052 Skhirat
16: 48569 Shimon_Peres
17: 19291 Morocco
18: 48569 Shimon_Peres
1: 19291 Morocco
    [pem_pos=0, final_scores=0.728953,
final_local_scores=0.74637836, pem=-0.2674848,
similarity=27.817757]
2: 9282173 Israel
    [pem_pos=0, final_scores=0.6681437,
final_local_scores=0.67217404, pem=-0.21279564,
similarity=25.305943]
3: 48569 Shimon_Peres
    [pem_pos=0, final_scores=0.71103525,
final_local_scores=0.7137309, pem=0.0, similarity=25.388866]
4: 4512052 Skhirat
    [pem_pos=0, final_scores=0.36870295,
final_local_scores=0.48296547, pem=0.0, similarity=18.502365]
5: 19291 Morocco
    [pem_pos=0, final_scores=0.6010308,
final_local_scores=0.6066428, pem=-0.29598162,
similarity=23.802736]
6: 410355 Hassan_II_of_Morocco
    [pem_pos=0, final_scores=0.39021719,
final_local_scores=0.38242579, pem=0.0, similarity=15.502062]
7: 19291 Morocco
    [pem_pos=0, final_scores=0.61716783,
final_local_scores=0.62427, pem=-0.29598162,
similarity=24.32877]
8: 9282173 Israel
    [pem_pos=0, final_scores=0.59062624,
final_local_scores=0.587498, pem=-0.34790626,
similarity=23.513792]
9: 48569 Shimon_Peres
    [pem_pos=0, final_scores=0.6191992,
final_local_scores=0.6060865, pem=0.0, similarity=22.176542]
10: 4512052 Skhirat
    [pem_pos=0, final_scores=0.38261783,
final_local_scores=0.49816537, pem=0.0, similarity=18.955961]
11: 57400 Rabat
    [pem_pos=0, final_scores=0.36433417,
final_local_scores=0.3671704, pem=-0.077565834,
similarity=15.468624]
```

```
12: 48569 Shimon_Peres
    [pem_pos=0, final_scores=0.5743447,
final_local_scores=0.55708975, pem=0.0, similarity=20.714382]
13: 19291 Morocco
    [pem_pos=0, final_scores=0.5487131,
final_local_scores=0.5494936, pem=-0.29598162,
similarity=22.097292]
14: 410355 Hassan_II_of_Morocco
    [pem_pos=0, final_scores=0.382374,
final_local_scores=0.37385827, pem=0.0, similarity=15.24639]
15: 4512052 Skhirat
    [pem_pos=0, final_scores=0.42329818,
final_local_scores=0.5426025, pem=0.0, similarity=20.282053]
16: 48569 Shimon_Peres
    [pem_pos=0, final_scores=0.7739759,
final_local_scores=0.78248394, pem=0.0, similarity=27.440594]
17: 19291 Morocco
    [pem_pos=0, final_scores=0.58394694,
final_local_scores=0.5879813, pem=-0.29598162,
similarity=23.24584]
18: 48569 Shimon_Peres
    [pem_pos=0, final_scores=0.6593747,
final_local_scores=0.6572996, pem=0.0, similarity=23.704844]
```

Sample: b'860_Scandal'

Scandal hits [1 1 Clinton 1] campaign at vital moment . Michael Conlon [2 2 CHICAGO 2] 1996-08-29 President [3 3 Bill Clinton 3] 's triumphal appearance at the [4 4 Democratic 4] convention , a vital moment in his bid for a second term , was marred on Thursday by the resignation of a top adviser in a reported sex scandal . [5 5 Clinton 5] was at work on the nomination acceptance speech that will launch his 10-week re-election campaign when political strategist [6 6 Dick Morris 6] abruptly announced his resignation on Thursday . The tabloid Star magazine reported the married [7 7 Morris 7] had a lengthy affair with a \$ 200-an-hour prostitute who he allowed to eavesdrop on telephone conversations with [8 8 Clinton 8] , and with whom he shared [9 9 White House 9] speeches before they were made . " [10 10 Dick Morris 10] is my friend , and he is a superb political strategist , " [11 11 Clinton 11] said in a written statement . " I am and always will be grateful for the great contributions he has made to my campaigns and for the invaluable work he has done for me over the last two years . " [12 12 Morris 12] declined to address the allegations directly in his resignation statement but said he was quitting to avoid becoming a campaign issue . The surprise development captivated the thousands of reporters at the convention and clearly worried some [13 13 Democrats 13] , who had planned a triumphal celebration of [14 14 Clinton 14] 's lead over [15 15 Dole 15] in the opinion polls . Senator [16 16 Dianne Feinstein 16] , a [17 17 California 17] [18 18 Democrat 18] , called it a " big bump " in

the way of the [19 19 Clinton 19] campaign . " It comes at the worst possible time on one of the biggest days for the president , " [20 20 Feinstein 20] said . [21 21 Clinton 21] was to deliver his acceptance speech at the final session of the [22 22 Democratic Convention 22] opening at 8 p.m. EDT (midnight [23 23 GMT 23]) . The 50-year-old president has been dogged for years by allegations of financial wrongdoing , sexual misconduct and questionable judgment in selecting his advisers . [24 24 Republicans 24] hope to seize on the " character " issue to bolster [25 25 Clinton 25] 's challenger [26 26 Bob Dole 26] in the final weeks of the campaign . Speaking in [27 27 Santa Barbara 27] , [28 28 California 28] , [29 29 Dole 29] did not directly refer to the sexual scandal but said the departure of [30 30 Morris 30] , who had advised [31 31 Clinton 31] to chart a more centrist political course , would make [32 32 Clinton 32] drift back to the left . " [33 33 Morris 33] has been trying to make President [34 34 Clinton 34] a [35 35 Republican 35] , now maybe he 'll revert to the liberal [36 36 Democrat 36] that he really is , " [37 37 Dole 37] told reporters . [38 38 Clinton 38] will hit the road on Friday for a bus tour across several states in his fight to become the first [39 39 Democratic 39] incumbent re-elected to a second term since the days of [40 40 Franklin D. Roosevelt 40] . Aides said [41 41 Clinton 41] planned to spend most of Thursday in his hotel room several miles (km) from the convention hall working on his speech . It was apparently neglected on his long " whistle-stop " train trip to the convention , during which he revelled in contact with friendly crowds across the country 's heartland . He was also hoarse and was giving his voice a rest . But first lady [42 42 Hillary Rodham Clinton 42] told [43 43 ABC 43] television , " He 's really fired up . " " He wants to outline to the [44 44 American 44] people what he thinks has been accomplished in the last four years , and what he would like to see done in the next four years , " she said . " He 's very excited about this convention ... He 's excited about the campaign . But more than that , he 's very resolute about what he wants to do . " " I do n't take anything for granted . I always expect elections to get close . I expect to have a great deal of up and down days between now and then , " she said of the November 5 election date . [45 45 Clinton 45] is leading [46 46 Dole 46] by as much as 15 points according to some polls . The campaign pits [47 47 Dole 47] , a man of quick and withering wit and long-time public service , but stilted speaking style , against the glib and confident [48 48 Clinton 48] , who has perfected a style that makes direct eye contact with his audience .

1: 3356 Bill_Clinton

2: 6886 Chicago

3: 3356 Bill_Clinton

4: 5043544 Democratic_Party_(United_States)

5: 3356 Bill_Clinton

6: 1068374 Dick_Morris
7: 1068374 Dick_Morris
8: 3356 Bill_Clinton
9: 33057 White_House
10: 1068374 Dick_Morris
11: 3356 Bill_Clinton
12: 1068374 Dick_Morris
13: 5043544 Democratic_Party_(United_States)
14: 3356 Bill_Clinton
15: 55846 Bob_Dole
16: 224672 Dianne_Feinstein
17: 5407 California
18: 5043544 Democratic_Party_(United_States)
19: 3356 Bill_Clinton
20: 224672 Dianne_Feinstein
21: 3356 Bill_Clinton
22: 468102 Democratic_National_Convention
23: 12701 Greenwich_Mean_Time
24: 32070 Republican_Party_(United_States)
25: 3356 Bill_Clinton
26: 55846 Bob_Dole
27: 108139 Santa_Barbara,_California
28: 5407 California
29: 55846 Bob_Dole
30: 1068374 Dick_Morris
31: 3356 Bill_Clinton
32: 3356 Bill_Clinton
33: 1068374 Dick_Morris
34: 3356 Bill_Clinton
35: 32070 Republican_Party_(United_States)
36: 5043544 Democratic_Party_(United_States)
37: 55846 Bob_Dole
38: 3356 Bill_Clinton
39: 5043544 Democratic_Party_(United_States)
40: 10979 Franklin_D._Roosevelt
41: 3356 Bill_Clinton
42: 5043192 Hillary_Rodham_Clinton
43: 62027 American_Broadcasting_Company
44: 3434750 United_States
45: 3356 Bill_Clinton
46: 55846 Bob_Dole
47: 55846 Bob_Dole
48: 3356 Bill_Clinton
1: 3356 Bill_Clinton
[pem_pos=1, final_scores=0.84275943,
final_local_scores=0.80677325, pem=0.0, similarity=28.165436]
2: 6886 Chicago
[pem_pos=0, final_scores=0.5901155,
final_local_scores=0.63451236, pem=-0.20658566,
similarity=24.148273]
3: 3356 Bill_Clinton
[pem_pos=0, final_scores=0.6482046,

```
final_local_scores=0.5937843, pem=0.0, similarity=21.80942]
4: 5043544 Democratic_Party_(United_States)
    [pem_pos=1, final_scores=0.46784872,
final_local_scores=0.44734365, pem=-0.076093495,
similarity=17.85315]
5: 3356 Bill_Clinton
    [pem_pos=1, final_scores=0.89045316,
final_local_scores=0.8588713, pem=0.0, similarity=29.720144]
6: 1068374 Dick_Morris
    [pem_pos=0, final_scores=0.684745,
final_local_scores=0.7212159, pem=0.0, similarity=25.612234]
7: 1068374 Dick_Morris
    [pem_pos=0, final_scores=0.8465913,
final_local_scores=0.8980084, pem=0.0, similarity=30.888075]
8: 3356 Bill_Clinton
    [pem_pos=1, final_scores=0.890919,
final_local_scores=0.8593802, pem=0.0, similarity=29.735332]
9: 33057 White_House
    [pem_pos=0, final_scores=0.47009832,
final_local_scores=0.44564086, pem=-0.0056634876,
similarity=17.419323]
10: 1068374 Dick_Morris
    [pem_pos=0, final_scores=0.64917237,
final_local_scores=0.68235815, pem=0.0, similarity=24.45264]
11: 3356 Bill_Clinton
    [pem_pos=1, final_scores=0.84792787,
final_local_scores=0.81241894, pem=0.0, similarity=28.333914]
12: 1068374 Dick_Morris
    [pem_pos=0, final_scores=0.78447825,
final_local_scores=0.83015937, pem=0.0, similarity=28.863325]
13: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.4740479,
final_local_scores=0.45411533, pem=-0.18007085,
similarity=18.620674]
14: 3356 Bill_Clinton
    [pem_pos=1, final_scores=0.8482984,
final_local_scores=0.81282365, pem=0.0, similarity=28.345991]
15: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.8655997,
final_local_scores=0.86657375, pem=0.0, similarity=29.95]
16: 224672 Dianne_Feinstein
    [pem_pos=0, final_scores=0.6743968,
final_local_scores=0.6853039, pem=0.0, similarity=24.540548]
17: 5407 California
    [pem_pos=0, final_scores=0.5919731,
final_local_scores=0.6549439, pem=-0.08662026,
similarity=24.1056]
18: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.3986082,
final_local_scores=0.37170893, pem=-0.061652232,
similarity=15.517524]
19: 3356 Bill_Clinton
```

```
[pem_pos=1, final_scores=0.8913768,
final_local_scores=0.85988027, pem=0.0, similarity=29.750256]
20: 224672 Dianne_Feinstein
    [pem_pos=0, final_scores=0.58502287,
final_local_scores=0.5900494, pem=0.0, similarity=21.697964]
21: 3356 Bill_Clinton
    [pem_pos=1, final_scores=0.8933422,
final_local_scores=0.86202717, pem=0.0, similarity=29.814323]
22: 468102 Democratic_National_Convention
    [pem_pos=0, final_scores=0.39358497,
final_local_scores=0.37523097, pem=-0.23190276,
similarity=16.548477]
23: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.53585994,
final_local_scores=0.6886827, pem=0.0, similarity=24.641378]
24: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.40783602,
final_local_scores=0.3810749, pem=-0.11427376,
similarity=16.083187]
25: 3356 Bill_Clinton
    [pem_pos=1, final_scores=0.8680962,
final_local_scores=0.8344497, pem=0.0, similarity=28.991356]
26: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.79331666,
final_local_scores=0.7860796, pem=0.0, similarity=27.547895]
27: 108139 Santa_Barbara,_California
    [pem_pos=0, final_scores=0.33415467,
final_local_scores=0.389153, pem=-0.4948885,
similarity=18.394094]
28: 5407 California
    [pem_pos=0, final_scores=0.6144975,
final_local_scores=0.6795483, pem=-0.08662026,
similarity=24.839846]
29: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.8112734,
final_local_scores=0.80723053, pem=0.0, similarity=28.179083]
30: 1068374 Dick_Morris
    [pem_pos=0, final_scores=0.84533626,
final_local_scores=0.8966375, pem=0.0, similarity=30.847166]
31: 3356 Bill_Clinton
    [pem_pos=1, final_scores=0.8278269,
final_local_scores=0.79046166, pem=0.0, similarity=27.678665]
32: 3356 Bill_Clinton
    [pem_pos=1, final_scores=0.8081973,
final_local_scores=0.76901937, pem=0.0, similarity=27.038784]
33: 1068374 Dick_Morris
    [pem_pos=0, final_scores=0.7598813,
final_local_scores=0.80329084, pem=0.0, similarity=28.061514]
34: 3356 Bill_Clinton
    [pem_pos=1, final_scores=0.8523665,
final_local_scores=0.8172675, pem=0.0, similarity=28.478605]
35: 32070 Republican_Party_(United_States)
```

```
[pem_pos=0, final_scores=0.40243232,
final_local_scores=0.37517214, pem=-0.1310191,
similarity=15.9981]
36: 5043544 Democratic_Party_(United_States)
[pem_pos=0, final_scores=0.46509367,
final_local_scores=0.44433415, pem=-0.061652232,
similarity=17.684805]
37: 55846 Bob_Dole
[pem_pos=0, final_scores=0.939097,
final_local_scores=0.94685835, pem=0.0, similarity=32.345856]
38: 3356 Bill_Clinton
[pem_pos=1, final_scores=0.8669329,
final_local_scores=0.83317906, pem=0.0, similarity=28.953438]
39: 5043544 Democratic_Party_(United_States)
[pem_pos=0, final_scores=0.4820121,
final_local_scores=0.46281505, pem=-0.076093495,
similarity=18.314846]
40: 10979 Franklin_D._Roosevelt
[pem_pos=0, final_scores=0.41168088,
final_local_scores=0.38219774, pem=0.0, similarity=15.495256]
41: 3356 Bill_Clinton
[pem_pos=1, final_scores=0.84191483,
final_local_scores=0.80585057, pem=0.0, similarity=28.137901]
42: 5043192 Hillary_Rodham_Clinton
[pem_pos=0, final_scores=0.5094564,
final_local_scores=0.4704122, pem=0.0, similarity=18.12775]
43: 62027 American_Broadcasting_Company
[pem_pos=0, final_scores=0.48911542,
final_local_scores=0.54948956, pem=-0.3336378,
similarity=22.301952]
44: 3434750 United_States
[pem_pos=0, final_scores=0.47664315,
final_local_scores=0.51158303, pem=-0.12994486,
similarity=20.063032]
45: 3356 Bill_Clinton
[pem_pos=1, final_scores=0.84965533,
final_local_scores=0.8143059, pem=0.0, similarity=28.390224]
46: 55846 Bob_Dole
[pem_pos=0, final_scores=0.8982368,
final_local_scores=0.9022247, pem=0.0, similarity=31.013899]
47: 55846 Bob_Dole
[pem_pos=0, final_scores=0.96521455,
final_local_scores=0.97538775, pem=0.0, similarity=33.197227]
48: 3356 Bill_Clinton
[pem_pos=1, final_scores=0.8874242,
final_local_scores=0.8555626, pem=0.0, similarity=29.621408]
```

Sample: b'861_Montana'

```
[1 1 Montana 1] weekly muni bond indices - [2 2 Piper Jaffray
2] . [3 3 NEW YORK 3] 1996-08-29 The following [4 4 Montana 4]
tax-exempt municipal bond indices were compiled by Piper
Jaffray Inc for the week ending August 30 . Previous 8/30
```

Week Change ----- Year A-rated Gen'l
Obligation 4.45 % 4.40 % +0.05 10 Year A-rated Gen'l
Obligation 4.90 % 4.90 % ----- 15 Year A-rated Gen'l
Obligation 5.40 % 5.35 % +0.05 20 Year A-rated Gen'l
Obligation 5.55 % 5.50 % +0.05 30 Year A-rated Housing Rev
6.05 % 6.00 % +0.05 -- U.S. Municipal Desk , 212-859-1650

1: 19978 Montana
2: 2136403 Piper_Jaffray
3: 645042 New_York_City
4: 19978 Montana
1: 19978 Montana
[pem_pos=0, final_scores=0.5164851,
final_local_scores=0.50571537, pem=-0.308461,
similarity=20.858727]
2: 2136403 Piper_Jaffray
[pem_pos=0, final_scores=0.46892864,
final_local_scores=0.6281528, pem=0.0, similarity=22.835045]
3: 645042 New_York_City
[pem_pos=1, final_scores=0.404603,
final_local_scores=0.44512314, pem=-1.7906687,
similarity=27.11101]
4: 19978 Montana
[pem_pos=0, final_scores=0.3438627,
final_local_scores=0.31715167, pem=-0.308461,
similarity=15.23161]

Sample: b'862_Researchers'

Researchers report progress in muscular dystrophy . [1 1 PHI
LADELPHIA 1] 1996-08-29 [2 2 University of Pennsylvania 2]
researchers on Thursday said a new gene-therapy technique for
treating muscular dystrophy disease had shown progress in
laboratory animals . Word of the findings , to be published
in the Oct. 1 issue of the journal " Human Gene Therapy , "
came in advance of the annual Jerry Lewis Labour Day weekend
telethon to raise money for muscular dystrophy research .
Several hurdles must be overcome before the method is used in
human trials . Nevertheless , " a treatment based on the new
strategy may have the potential to benefit many
patients , " the [3 3 University of Pennsylvania 3] Medical
Centre said in a release . Muscular dystrophy is a fatal
illness in which the body 's muscle tissue degenerates and is
replaced by fat . Death strikes in early adulthood .
Individuals with the disease have a non-working version of a
gene responsible for producing a crucial muscle protein
called dystrophin . In the study at the University 's
Institute for Human Gene Therapy , researchers altered a
common-cold virus to carry a version of the working
dystrophin gene . The virus , which also was altered to
minimise its susceptibility to the immune system , was then
injected into the muscle cells of mice bred to lack
dystrophin genes . In the experiment , between 30 to 40
percent of the muscle fibers in one group of mice produced

dystrophin for two weeks before diminishing . Similar results had been obtained previously in test-tube cell cultures , but not in live animals , the university said . The university said methods are still needed to make enough of the altered virus to treat humans , to further decrease the immune-system response to the virus , and to deliver the virus to human muscle tissue .

```
1: 50585 Philadelphia
2: 31793 University_of_Pennsylvania
3: 31793 University_of_Pennsylvania
1: 50585 Philadelphia
    [pem_pos=0, final_scores=0.59539276,
final_local_scores=0.6104682, pem=-0.2107264,
similarity=23.453264]
2: 31793 University_of_Pennsylvania
    [pem_pos=0, final_scores=0.49864757,
final_local_scores=0.43929344, pem=0.0, similarity=17.199106]
3: 31793 University_of_Pennsylvania
    [pem_pos=0, final_scores=0.48626077,
final_local_scores=0.42576277, pem=0.0, similarity=16.795322]
```

Sample: b'863_Export'

Export Business - Grain / oilseeds complex . [1 1 CHICAGO 1] 1996-08-29 Grain and oilseed exports reported by [2 2 USDA 2] and private export sources . WHEAT SALES - Taiwan Flour Mills Assn bought 98,000 tonnes of [3 3 U.S. 3] No.1 or No.2 wheat from [4 4 Cargill 4] , [5 5 Mitsui 5] , Continental and Louis Dreyfus Corp for shipment from the Pacific Northwest . For Sept 10-30 : 16,300 tonnes of dark northern spring (DNS) at \$ 212.00 ; 7,000 of hard red winter (HRW) at \$ 205.10 ; 2,700 of western white (WW) at \$ 202.65 . For Sept 20 - Oct 10 : 19,500 of DNS at \$ 212.25 , 10,000 of HRW at \$ 204.74 , 4,500 of WW at \$ 199.71 . For Sept 25 - Oct 20 : 23,500 of DNS at \$ 212.25 , 9,600 HRW at \$ 204.74 , 4,900 WW at \$ 199.56 . WHEAT SALES - The Commodity Credit Corp of USDA bought 18,278 tonnes of dark northern spring (DNS) wheat from Cargill Inc at \$ 195.79 per tonne , FOB , for donation to [6 6 Nicaragua 6] . Shipment is for Nov 15 - Dec 10 . WHEAT / BARLEY SALE - The Japanese Food Agency said it bought 20,000 tonnes of [7 7 U.S. 7] dark northern spring wheat , 20,000 of [8 8 Canadian 8] western red spring wheat , 20,000 of [9 9 Australian 9] standard white wheat and 20,000 of [10 10 Australian 10] feed barley at its weekly tender , all for October shipment . SOYBEAN SALES - The Taichung division of [1 11 Taiwan 11] 's Breakfast Soybean Procurement Assn bought 108,000 tonnes of [12 12 U.S. 12] soybeans . [13 13 Bunge 13] sold the first shipment and [14 14 Cargill 14] the second , traders said . For Nov 11-25 from the U.S. Gulf or Nov 26 - Dec 10 from PNW it paid \$ 0.8584 per bushel over CBOT January soybeans and for Dec 6-20 or Dec 21 - Jan it paid \$.8787 over CBOT January . - [15 15 Pakistan 15] bought 31,412 tonnes of PL-480 No.2 yellow soybeans from Cargill Inc for \$

303.19 per tonne , FOB U.S. Gulf , agents for the buyer said . The soybeans were for Oct 15-30 shipment . BARLEY TENDER - The Cyprus Grain Commission said it invited offers September 3 to supply 25,000 tonnes of feed barley , with shipment for Sept 25 - Oct 10 from [16 16 Europe 16] or Sept 15-30 from [17 17 North America 17] . MARKET TALK - [18 18 Sri Lanka 18] plans to import up to 400,000 tonnes of rice by the end of this year to meet a crop shortfall caused by drought and rising demand , government officials said on Thursday . MARKET TALK - [19 19 USDA 19] net change in weekly export commitments for the week ended August 22 , includes old crop and new crop , were : wheat up 595,400 tonnes old , nil new ; corn up 1,900 old , up 319,600 new ; soybeans down 12,300 old , up 300,800 new ; upland cotton up 50,400 bales new , nil old ; soymeal 54,800 old , up 100,600 new , soyoil nil old , up 75,000 new ; barley up 1,700 old , nil new ; sorghum 6,200 old , up 156,700 new ; pima cotton up 4,000 bales old , nil new ; rice up 49,900 old , nil new ... [20 20 USDA 20] Thursday forecast [21 21 U.S. 21] agricultural exports in fiscal year 1997 would decline to \$ 58 billion , down \$ 2 billion from the record \$ 60 billion seen for fiscal 1996 . Oilseed exports are expected to rise by \$ 800 million and livestock , poultry and fruits and vegetables are seen gaining more than \$ 1 billion . [22 22 USDA 22] pegged [23 23 U.S. 23] wheat exports at 25.0 million tonnes in fiscal 1997 versus 32.0 million tonnes the prior year ... The [24 24 European Union 24] agreed on Thursday to increase by 300,000 tonnes the quota of [25 25 German 25] intervention barley available for export , [26 26 France 26] ONIC said . The EU 's grain panel will add two tranches of 150,000 tonnes each to the existing allocation , it said . -- [27 27 Chicago 27] newsdesk 312-408-8720--

1: 6886 Chicago
2: 70896 United_States_Department_of_Agriculture
3: 3434750 United_States
4: 952282 Cargill
5: 290271 Mitsui
6: 21362 Nicaragua
7: 3434750 United_States
8: 5042916 Canada
9: 4689264 Australia
10: 4689264 Australia
11: 25734 Taiwan
12: 3434750 United_States
13: 54998 Bunge_Limited
14: 952282 Cargill
15: 23235 Pakistan
16: 9239 Europe
17: 21139 North_America
18: 26750 Sri_Lanka
19: 70896 United_States_Department_of_Agriculture
20: 70896 United_States_Department_of_Agriculture

```
21: 3434750 United_States
22: 70896 United_States_Department_of_Agriculture
23: 3434750 United_States
24: 9317 European_Union
25: 11867 Germany
26: 5843419 France
27: 6886 Chicago
1: 6886 Chicago
    [pem_pos=0, final_scores=0.6210611,
final_local_scores=0.63372415, pem=-0.20658566,
similarity=24.124752]
2: 70896 United_States_Department_of_Agriculture
    [pem_pos=0, final_scores=0.39042807,
final_local_scores=0.4037164, pem=0.0, similarity=16.137415]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.5571643,
final_local_scores=0.5108337, pem=-0.14075686,
similarity=20.099466]
4: 952282 Cargill
    [pem_pos=0, final_scores=0.61714804,
final_local_scores=0.69812346, pem=0.0, similarity=24.92311]
5: 290271 Mitsui
    [pem_pos=0, final_scores=0.32220364,
final_local_scores=0.41466367, pem=0.0, similarity=16.464104]
6: 21362 Nicaragua
    [pem_pos=0, final_scores=0.5282796,
final_local_scores=0.5408917, pem=-0.24729352,
similarity=21.57582]
7: 3434750 United_States
    [pem_pos=0, final_scores=0.55902326,
final_local_scores=0.5128643, pem=-0.14075686,
similarity=20.160063]
8: 5042916 Canada
    [pem_pos=0, final_scores=0.641666,
final_local_scores=0.63448375, pem=-0.41303828,
similarity=25.27014]
9: 4689264 Australia
    [pem_pos=0, final_scores=0.5924341,
final_local_scores=0.5782494, pem=-0.491009,
similarity=24.01601]
10: 4689264 Australia
    [pem_pos=0, final_scores=0.5710867,
final_local_scores=0.5549306, pem=-0.491009,
similarity=23.320133]
11: 25734 Taiwan
    [pem_pos=0, final_scores=0.38195008,
final_local_scores=0.36834246, pem=-0.077886656,
similarity=15.505347]
12: 3434750 United_States
    [pem_pos=0, final_scores=0.54645044,
final_local_scores=0.49913043, pem=-0.14075686,
similarity=19.75022]
```

```
13: 54998 Bunge_Limited
    [pem_pos=2, final_scores=0.44211966,
final_local_scores=0.5332486, pem=-1.1289766,
similarity=26.142466]
14: 952282 Cargill
    [pem_pos=0, final_scores=0.62352806,
final_local_scores=0.70509267, pem=0.0, similarity=25.131084]
15: 23235 Pakistan
    [pem_pos=0, final_scores=0.5515706,
final_local_scores=0.5687538, pem=-0.21944217,
similarity=22.255821]
16: 9239 Europe
    [pem_pos=0, final_scores=0.53170437,
final_local_scores=0.5215, pem=-0.086338006,
similarity=20.121834]
17: 21139 North_America
    [pem_pos=0, final_scores=0.36614925,
final_local_scores=0.34939486, pem=-0.011697312,
similarity=14.579964]
18: 26750 Sri_Lanka
    [pem_pos=0, final_scores=0.4456113,
final_local_scores=0.4561047, pem=-0.3027737,
similarity=19.347317]
19: 70896 United_States_Department_of_Agriculture
    [pem_pos=0, final_scores=0.50276536,
final_local_scores=0.52642787, pem=0.0, similarity=19.79937]
20: 70896 United_States_Department_of_Agriculture
    [pem_pos=0, final_scores=0.583891,
final_local_scores=0.61504525, pem=0.0, similarity=22.44389]
21: 3434750 United_States
    [pem_pos=0, final_scores=0.58208245,
final_local_scores=0.538053, pem=-0.14075686,
similarity=20.911743]
22: 70896 United_States_Department_of_Agriculture
    [pem_pos=0, final_scores=0.51517373,
final_local_scores=0.5399821, pem=0.0, similarity=20.203856]
23: 3434750 United_States
    [pem_pos=0, final_scores=0.59493905,
final_local_scores=0.5520969, pem=-0.14075686,
similarity=21.330843]
24: 9317 European_Union
    [pem_pos=0, final_scores=0.43215817,
final_local_scores=0.4501196, pem=-0.005814796,
similarity=17.553802]
25: 11867 Germany
    [pem_pos=1, final_scores=0.68730897,
final_local_scores=0.70820194, pem=-1.0465823,
similarity=30.91535]
26: 5843419 France
    [pem_pos=0, final_scores=0.5034023,
final_local_scores=0.503432, pem=-0.16240834,
similarity=19.996328]
```

```
27: 6886 Chicago
    [pem_pos=0, final_scores=0.6994324,
final_local_scores=0.7193329, pem=-0.20658566,
similarity=26.679487]

Sample: b'864_USDA'
[1 1 USDA 1] gross cutout hide and offal value . DES MOINES
1996-08-29 The hide and offal value from a typical slaughter
steer for Thursday was estimated at $ 9.76 per cwt live , up
$ 0.03 when compared with Wednesday 's value . - [2 2 USDA 2]
1: 70896 United_States_Department_of_Agriculture
2: 70896 United_States_Department_of_Agriculture
1: 70896 United_States_Department_of_Agriculture
    [pem_pos=0, final_scores=0.56497294,
final_local_scores=0.5391635, pem=0.0, similarity=20.179426]
2: 70896 United_States_Department_of_Agriculture
    [pem_pos=0, final_scores=0.5027496,
final_local_scores=0.4711939, pem=0.0, similarity=18.15108]
```

```
Sample: b'865_Help-wanted'
Help-wanted ad index fell in July . [1 1 NEW YORK 1]
1996-08-29 The help-wanted advertising index fell in July ,
the [2 2 Conference Board 2] said Thursday , reflecting the
uneven nature of the nation 's labour markets . The monthly
index fell to 83.0 in July against a reading of 85.0 in
June , the private business research group said . In July ,
the volume of help-wanted advertising fell in five of the
nine [3 3 U.S. 3] regions . " The labour market has been
expanding throughout 1996 , but in a very uneven pattern , " [4 4 Conference Board 4] economist Ken Goldstein said . "
Recent want-ad figures indicate that conservative hiring
plans are keeping job growth below the rate of overall
economic activity . " With 2.5 percent gross domestic product
growth expected for 1996 , new job growth should slowly lower
the unemployment rate over the rest of the year . " With the
unemployment rate staying close to about 5.5 percent over the
last two years , there is a good chance the rate will slowly
drop to about 5.0 percent by the end of the year , "
Goldstein said . The July index matched the reading for
July , 1995 . The greatest declines in the volume of help-
wanted advertising were in the [5 5 New England 5] , Mountain
and [6 6 West South Central 6] regions . The greatest
increase was in the [7 7 East North Central 7] region .
1: 645042 New_York_City
2: 4338197 The_Conference_Board
3: 3434750 United_States
4: 4338197 The_Conference_Board
5: 21531764 New_England
6: 920048 West_South_Central_States
7: 920073 East_North_Central_States
1: 645042 New_York_City
    [pem_pos=1, final_scores=0.3731243,
```

```
final_local_scores=0.3840257, pem=-1.7906687,
similarity=25.28774]
2: 4338197 The_Conference_Board
    [pem_pos=0, final_scores=0.43241912,
final_local_scores=0.47608793, pem=0.0, similarity=18.297127]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.4974981,
final_local_scores=0.5061124, pem=-0.14075686,
similarity=19.958572]
4: 4338197 The_Conference_Board
    [pem_pos=0, final_scores=0.4820786,
final_local_scores=0.53033334, pem=0.0, similarity=19.915918]
5: 21531764 New_England
    [pem_pos=0, final_scores=0.45073277,
final_local_scores=0.4728604, pem=-0.13270761,
similarity=18.922495]
6: 920048 West_South_Central_States
    [pem_pos=0, final_scores=0.46091115,
final_local_scores=0.57203484, pem=0.0, similarity=21.160374]
7: 920073 East_North_Central_States
    [pem_pos=0, final_scores=0.35338843,
final_local_scores=0.45152128, pem=0.0, similarity=17.564009]
```

Sample: b'866_Police'

Police seek suspects in [1 1 Atlantic City 1] jewel heist . [2 2 ATLANTIC CITY 2] , [3 3 N.J. 3] 1996-08-29 [4 4 Atlantic City 4] police said Thursday they were seeking two men and two women in connection with a \$ 690,000 theft of jewelry and cash from a guest of the Showboat Hotel and Casino . Capt . Richard Andrews said police were seeking a man shown on a hotel videotape carrying a suitcase resembling the victim 's . " We want to talk to him , " Andrews said of the man . A second man and two women also were being sought . The thefts occurred Sunday when the victim , [5 5 New York 5] jewelry wholesaler Jerry Schein , left three suitcases in a closet at Somerset Jewellers in the hotel while he checked out . While he was gone , two women in their mid-twenties and an older man entered the jewelry store and tried to distract store owner Charles McGilley . When Schein returned two of the suitcases were missing . They contained \$ 650,000 in jewelry and \$ 40,000 in cash , Andrews said . He said the man on the videotape did not match the description given by the jewelry store owner .

```
1: 106211 Atlantic_City,_New_Jersey
2: 106211 Atlantic_City,_New_Jersey
3: 21648 New_Jersey
4: 106211 Atlantic_City,_New_Jersey
5: 645042 New_York_City
1: 106211 Atlantic_City,_New_Jersey
    [pem_pos=0, final_scores=0.38641047,
final_local_scores=0.312596, pem=-0.07798163,
similarity=13.842278]
```

```
2: 106211 Atlantic_City,_New_Jersey
    [pem_pos=0, final_scores=0.5407408,
final_local_scores=0.4811784, pem=-0.07798163,
similarity=18.873112]
3: 21648 New_Jersey
    [pem_pos=0, final_scores=0.62528265,
final_local_scores=0.61891997, pem=0.0, similarity=22.559519]
4: 106211 Atlantic_City,_New_Jersey
    [pem_pos=0, final_scores=0.39729595,
final_local_scores=0.32448673, pem=-0.07798163,
similarity=14.197121]
5: 645042 New_York_City
    [pem_pos=1, final_scores=0.4734854,
final_local_scores=0.4581023, pem=-1.7906687,
similarity=27.498333]
```

Sample: b'867_0.J.'

[1 1 0.J. Simpson 1] hints at more supporting evidence .
Jackie Frank [2 2 WASHINGTON 2] 1996-08-29 [3 3 0.J. Simpson 3]
] said on Thursday he was financially broken by his defence
against murder charges but he was hopeful new evidence to
support him would be available for a civil trial next month .
The former football star was found not guilty by a criminal
trial jury last October of the murders of his former wife , [4
4 Nicole Brown Simpson 4] , and her friend , [5 5 Ronald Gold
man 5] , in June 1994 . He now faces a civil suit brought by
families of the victims who hold him responsible for the
deaths . He told reporters a court order not to talk about
his case kept him from detailing how he has fulfilled his
pledge to find the killers . But he added without
elaborating , " hopefully we will see some things come out in
this next trial . " The judge in the civil trial has imposed
a sweeping gag order that prohibits lawyers , witnesses and
parties to the case from discussing it with the media or
elsewhere in public . " I would love to speak about
everything , " said [6 6 Simpson 6] , who vowed after his
acquittal to find the killers and offered a substantial
reward . His lawyers have said his defence in the civil trial
that starts Sept . 17 will be that he did not kill the
victims . [7 7 Simpson 7] said at the hotel news conference
his plans include eventually writing another book . He added
that he has had job offers but more have come from abroad
than at home . " I 'm broke . I am not crying the blues . I
can get along just fine , " he said . " Whatever you want to
send me , I need . " He again accused the news media of
erroneous reporting on his case but did not signal any plans
for lawsuits as he did on Wednesday in an address at a jam-
packed [8 8 Washington 8] church . Contrary to news reports ,
[9 9 Simpson 9] said , he has received support from both
blacks and whites . He again dismissed charges that he had
distanced himself from the black community during his
successful football and commercial career , only to seek

their support after he faced murder charges . A crowd of 2,000 paid \$ 10 a head to hear the former star running back for the [10 10 Buffalo Bills 10] professional football club and a [11 11 Football Hall of Fame 11] member on Wednesday night . The crowd in the church was wildly supportive , showering [12 12 Simpson 12] with gifts and praise . But outside , dozens of protesters from the D.C. Coalition Against Domestic Violence called for the church to support victims of violence instead .

1: 8955445 O._J._Simpson
2: 108956 Washington,_D.C.
3: 8955445 O._J._Simpson
4: 430434 Nicole_Brown_Simpson
5: 430438 Ronald_Goldman
6: 8955445 O._J._Simpson
7: 8955445 O._J._Simpson
8: 108956 Washington,_D.C.
9: 8955445 O._J._Simpson
10: 4315 Buffalo_Bills
11: 272590 Pro_Football_Hall_of_Fame
12: 8955445 O._J._Simpson
1: 8955445 O._J._Simpson
 [pem_pos=0, final_scores=0.4762146,
final_local_scores=0.40232772, pem=0.0, similarity=16.095974]
2: 108956 Washington,_D.C.
 [pem_pos=0, final_scores=0.5375234,
final_local_scores=0.6213185, pem=-1.7496481,
similarity=32.145954]
3: 8955445 O._J._Simpson
 [pem_pos=0, final_scores=0.5120116,
final_local_scores=0.44143057, pem=0.0, similarity=17.262882]
4: 430434 Nicole_Brown_Simpson
 [pem_pos=0, final_scores=0.3507856,
final_local_scores=0.40694696, pem=0.0, similarity=16.233822]
5: 430438 Ronald_Goldman
 [pem_pos=0, final_scores=0.48673564,
final_local_scores=0.5214318, pem=0.0, similarity=19.650278]
6: 8955445 O._J._Simpson
 [pem_pos=0, final_scores=0.84574586,
final_local_scores=0.8077347, pem=0.0, similarity=28.194128]
7: 8955445 O._J._Simpson
 [pem_pos=0, final_scores=0.8341858,
final_local_scores=0.7951071, pem=0.0, similarity=27.817295]
8: 108956 Washington,_D.C.
 [pem_pos=0, final_scores=0.41361266,
final_local_scores=0.48596483, pem=-1.7496481,
similarity=28.106731]
9: 8955445 O._J._Simpson
 [pem_pos=0, final_scores=0.894968,
final_local_scores=0.86150247, pem=0.0, similarity=29.798664]
10: 4315 Buffalo_Bills
 [pem_pos=0, final_scores=0.29425806,

```
final_local_scores=0.31494862, pem=-0.11010485,
similarity=14.087175]
11: 272590 Pro_Football_Hall_of_Fame
    [pem_pos=0, final_scores=0.40751916,
final_local_scores=0.4280911, pem=0.0, similarity=16.864805]
12: 8955445 O._J._Simpson
    [pem_pos=0, final_scores=0.8511998,
final_local_scores=0.8136924, pem=0.0, similarity=28.371918]
```

Sample: b'868_Clinton'

[1 1 Clinton 1] adviser [2 2 Morris 2] announces resignation . [3 3 CHICAGO 3] 1996-08-29 President [4 4 Bill C linton 4] 's top political strategist [5 5 Dick Morris 5] resigned on Thursday , saying he did not want to become an issue in [6 6 Clinton 6] 's re-election campaign . In a written statement , distributed by the [7 7 Clinton 7] campaign , [8 8 Morris 8] avoided comment on published allegations that he had engaged in a year-long affair with a \$ 200-an-hour prostitute . The statement from [9 9 Morris 9] said that he had submitted his resignation on Wednesday night . " While I served I sought to avoid the limelight because I did not want to become the message . Now , I resign so I will not become the issue , " he said . The announcement followed a report in the weekly supermarket tabloid Star magazine , reprinted in Thursday 's editions of the [10 10 New York Post 10] , that the married adviser had hired a 37-year-old prostitute on a weekly basis while visiting [11 11 Washington 11] to advise [12 12 Clinton 12] on his re-election campaign . " I will not subject my wife , family or friends to the sadistic , vitriol of yellow journalism . I will not dignify such journalism with a reply or an answer . I never will , " his statement said . It was distributed to reporters at the press centre of [13 13 Clinton 13] 's [14 14 Democratic 14] convention headquarters just hours before the president was to address the delegates accepting the party 's nomination for a second four-year term in the [15 15 White Ho use 15] .

1: 3356 Bill_Clinton
2: 1068374 Dick_Morris
3: 6886 Chicago
4: 3356 Bill_Clinton
5: 1068374 Dick_Morris
6: 3356 Bill_Clinton
7: 3356 Bill_Clinton
8: 1068374 Dick_Morris
9: 1068374 Dick_Morris
10: 102227 New_York_Post
11: 108956 Washington,_D.C.
12: 3356 Bill_Clinton
13: 3356 Bill_Clinton
14: 5043544 Democratic_Party_(United_States)
15: 33057 White_House

```
1: 3356 Bill_Clinton
   [pem_pos=0, final_scores=0.7794687,
final_local_scores=0.73780364, pem=0.0, similarity=26.107244]
2: 1068374 Dick_Morris
   [pem_pos=0, final_scores=0.7586977,
final_local_scores=0.7924635, pem=0.0, similarity=27.738403]
3: 6886 Chicago
   [pem_pos=0, final_scores=0.59418786,
final_local_scores=0.6421793, pem=-0.20658566,
similarity=24.37707]
4: 3356 Bill_Clinton
   [pem_pos=0, final_scores=0.6198712,
final_local_scores=0.5618281, pem=0.0, similarity=20.855783]
5: 1068374 Dick_Morris
   [pem_pos=0, final_scores=0.7210365,
final_local_scores=0.7513244, pem=0.0, similarity=26.510733]
6: 3356 Bill_Clinton
   [pem_pos=0, final_scores=0.83594304,
final_local_scores=0.79949325, pem=0.0, similarity=27.948187]
7: 3356 Bill_Clinton
   [pem_pos=0, final_scores=0.8819435,
final_local_scores=0.8497418, pem=0.0, similarity=29.447704]
8: 1068374 Dick_Morris
   [pem_pos=0, final_scores=0.6868189,
final_local_scores=0.7139468, pem=0.0, similarity=25.39531]
9: 1068374 Dick_Morris
   [pem_pos=0, final_scores=0.69147915,
final_local_scores=0.7190375, pem=0.0, similarity=25.547226]
10: 102227 New_York_Post
    [pem_pos=0, final_scores=0.4100955,
final_local_scores=0.45013708, pem=0.0, similarity=17.522701]
11: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.46936047,
final_local_scores=0.46726072, pem=-1.7496481,
similarity=27.548563]
12: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.88163763,
final_local_scores=0.8494077, pem=0.0, similarity=29.437733]
13: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.9078714,
final_local_scores=0.8780641, pem=0.0, similarity=30.292896]
14: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.43904918,
final_local_scores=0.44356173, pem=-0.076093495,
similarity=17.740288]
15: 33057 White_House
    [pem_pos=0, final_scores=0.49842238,
final_local_scores=0.4708379, pem=-0.0056634876,
similarity=18.171253]
```

Sample: b'869_U.S.'

[1 1 U.S. 1] surgeon investigated for discarding foot . [2 2

CHARLESTON 2] , S.C. 1996-08-29 Health officials said on Thursday they were investigating the discovery of an amputated foot on a beach near [3 3 Charleston 3] to determine whether a local surgeon had improperly disposed of infectious waste . The foot , which washed up on Sullivan 's Island beach this month , was amputated three years ago from a child whose legs were deformed . The foot had to be removed so the infant could be fitted with a prosthesis . An orthopedic surgeon was given permission by the child 's parents and the hospital to keep the foot for research and educational purposes . Health officials said the surgeon told authorities he stored the foot in his freezer at home , but the freezer recently broke down and the contents spoiled . The surgeon , who apologised for the incident , said he decided to put the foot in a crab trap to remove the flesh . The foot later washed up on the beach .

1: 3434750 United_States

2: 61024 Charleston,_South_Carolina

3: 61024 Charleston,_South_Carolina

1: 3434750 United_States

[pem_pos=0, final_scores=0.43304664, final_local_scores=0.46126765, pem=-0.14075686, similarity=18.620317]

2: 61024 Charleston,_South_Carolina

[pem_pos=0, final_scores=0.35384864, final_local_scores=0.28865448, pem=-0.98667276, similarity=18.069418]

3: 61024 Charleston,_South_Carolina

[pem_pos=0, final_scores=0.30699527, final_local_scores=0.23747423, pem=-0.98667276, similarity=16.542097]

Sample: b'870_Florida'

Florida cop disguised as shrub nabs bad guys . [1 1 MIAMI 1] 1996-08-29 It was a bush that bagged the bad guys . When four would-be robbers , armed and masked , showed up to rob a Checker 's restaurant in the [2 2 Fort Lauderdale 2] suburb of [3 3 Pembroke Pines 3] on Tuesday , they had no idea the shrub near the drive-through window was toting a shotgun . Detective Earl Feugill , camouflaged as a shaggy green bush , ordered them to freeze . " They were quite surprised , " he told the [4 4 Miami Herald 4] . Feugill said he made the hot , heavy suit , which he first used in the [5 5 Marines 5] , by attaching strips of burlap to a camouflage outfit . Green and black face paint completed his disguise . He was staking out the restaurant after a series of robberies at local fast-food places . [6 6 Pembroke Pines 6] police said five people were arrested as a result of the 90-minute stakeout , including the four robbers and a restaurant employee who was allegedly prepared to let them in a back door .

1: 53846 Miami

```
2: 109028 Fort_Lauderdale,_Florida
3: 109061 Pembroke_Pines,_Florida
4: 607797 The_Miami_Herald
5: 17349325 United_States_Marine_Corps
6: 109061 Pembroke_Pines,_Florida
1: 53846 Miami
    [pem_pos=0, final_scores=0.54657716,
final_local_scores=0.52216864, pem=-0.35396606,
similarity=21.597189]
2: 109028 Fort_Lauderdale,_Florida
    [pem_pos=0, final_scores=0.40661734,
final_local_scores=0.3675412, pem=0.0, similarity=15.057876]
3: 109061 Pembroke_Pines,_Florida
    [pem_pos=0, final_scores=0.47304887,
final_local_scores=0.45599347, pem=0.0, similarity=17.697468]
4: 607797 The_Miami_Herald
    [pem_pos=0, final_scores=0.46852845,
final_local_scores=0.5123189, pem=0.0, similarity=19.37833]
5: 17349325 United_States_Marine_Corps
    [pem_pos=0, final_scores=0.2676633,
final_local_scores=0.3039324, pem=-0.45372075,
similarity=15.627066]
6: 109061 Pembroke_Pines,_Florida
    [pem_pos=0, final_scores=0.48349965,
final_local_scores=0.46740937, pem=0.0, similarity=18.038141]
```

Sample: b'871_U.S.'

U.S. DLA sets tin price at \$ 2.7975 per lb . [1 1 WASHINGTON
1] 1996-08-29 The U.S. Defense Logistics Agency set Thursday
's offering price for stockpile tin at \$ 2.7975 per lb ,
versus \$ 2.7775 per lb yesterday .

```
1: 108956 Washington,_D.C.
1: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.2833473,
final_local_scores=0.58210105, pem=-1.7496481,
similarity=30.975628]
```

Sample: b'872_Key'

Key [1 1 Clinton 1] aide resigns , [2 2 NBC 2] says . [3 3 CH
ICAGO 3] 1996-08-29 [4 4 Dick Morris 4] , the [5 5 Republican
5] political consultant who reshaped [6 6 U.S. 6] President [7
7 Bill Clinton 7] 's reelection campaign , has resigned , MS-
NBC News reported Thursday . [8 8 Morris 8] drew the ire of
liberal [9 9 Clinton 9] aides for repositioning the president
in the political centre . There was no immediate comment on
the report from the [10 10 White House 10] .

```
1: 3356 Bill_Clinton
2: 21780 NBC
3: 6886 Chicago
4: 1068374 Dick_Morris
5: 32070 Republican_Party_(United_States)
6: 3434750 United_States
```

```
7: 3356 Bill_Clinton
8: 1068374 Dick_Morris
9: 3356 Bill_Clinton
10: 33057 White_House
1: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.79401404,
final_local_scores=0.76711875, pem=0.0, similarity=26.982067]
2: 21780 NBC
    [pem_pos=0, final_scores=0.4821458,
final_local_scores=0.52170277, pem=-0.06403453,
similarity=20.006594]
3: 6886 Chicago
    [pem_pos=0, final_scores=0.62917477,
final_local_scores=0.6640344, pem=-0.20658566,
similarity=25.02927]
4: 1068374 Dick_Morris
    [pem_pos=0, final_scores=0.6402344,
final_local_scores=0.67889357, pem=0.0, similarity=24.34925]
5: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.44136745,
final_local_scores=0.44251335, pem=-0.1310191,
similarity=18.007696]
6: 3434750 United_States
    [pem_pos=0, final_scores=0.5352469,
final_local_scores=0.56375796, pem=-0.14075686,
similarity=21.678833]
7: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.56630903,
final_local_scores=0.5129285, pem=0.0, similarity=19.396523]
8: 1068374 Dick_Morris
    [pem_pos=0, final_scores=0.6883711,
final_local_scores=0.7314756, pem=0.0, similarity=25.918404]
9: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.9165661,
final_local_scores=0.9009883, pem=0.0, similarity=30.977]
10: 33057 White_House
    [pem_pos=0, final_scores=0.44578058,
final_local_scores=0.41485095, pem=-0.0056634876,
similarity=16.500492]
```

Sample: b'873_U.S.'

```
[1 1 U.S. 1] corn gluten meal steady-higher , feed flat . [2
2 CHICAGO 2] 1996-08-29 [3 3 U.S. 3] corn gluten feed prices
were flat while meal values were steady to firmer on
Thursday . Dealers noted a seasonal pickup in meal demand .
-- CHICAGO AREA MILLS ( dollars per short ton ) Gluten feed
21 pct bulk Spot - 117.00 unc Gluten feed pellets Spot - unq
Gluten meal 60 pct bulk rail Spot - 320.00 up 5 [4 4 DECATUR 4
] , IL / CLINTON AND CEDAR RAPIDS , IA Gluten feed 18 pct
pellets Spot - 117.00 unc Gluten meal 60 pct bulk Spot -
310.00 unc ( [5 5 Chicago 5] newsdesk 312-408-8720 )
1: 3434750 United_States
```

```
2: 6886 Chicago
3: 3434750 United_States
4: 111543 Decatur,_Illinois
5: 6886 Chicago
1: 3434750 United_States
    [pem_pos=0, final_scores=0.4942065,
final_local_scores=0.45416182, pem=-0.14075686,
similarity=18.408264]
2: 6886 Chicago
    [pem_pos=0, final_scores=0.65311545,
final_local_scores=0.6135922, pem=-0.20658566,
similarity=23.523973]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.54422545,
final_local_scores=0.50879997, pem=-0.14075686,
similarity=20.038778]
4: 111543 Decatur,_Illinois
    [pem_pos=1, final_scores=0.12649138,
final_local_scores=0.09839246, pem=-1.7115579,
similarity=16.333658]
5: 6886 Chicago
    [pem_pos=0, final_scores=0.7049827,
final_local_scores=0.6702493, pem=-0.20658566,
similarity=25.214735]
```

Sample: b'874_EU'

EU barley sale worth \$ 145 / T , for [1 1 Saudi 1] -
sources . [2 2 PARIS 2] 1996-08-29 A [3 3 European Union 3]
sale of 234,324 tonnes of [4 4 German 4] intervention barley
is worth some \$ 145 per tonne fob [5 5 Germany 5] and is
mostly destined for [6 6 Saudi Arabia 6] , [7 7 European 7]
grain sources said on Thursday . The EU 's cereals management
committee sold 234,324 tonnes of [8 8 German 8] intervention
barley at a minimum price of 105.07 [9 9 Ecus 9] per tonne . [
10 10 Saudi Arabia 10] provisionally bought 800,000 tonnes of
optional-origin barley at an August 21 tender at prices
between \$ 160 and \$ 162 including cost , insurance and
freight , traders said last week . But [11 11 European 11]
grain traders and officials said the [12 12 Saudis 12] might
reduce the purchase to 600,000 tonnes . Traders have said a
substantial part of the deal was likely to come from the [13 1
3 European Union 13] , which enjoys a supply and freight
advantage over other producers . But subtracting freight
costs , the equivalent fob price of the deal is around \$
142 , well below the \$ 149 per tonne floor price which the EU
put on its barley as news of the deal emerged last week .
Last Thursday the EU sold 34,277 tonnes of [14 14 German 14]
intervention barley at a minimum price of 109.36 [15 15 Ecus 1
5] per tonne , which was seen as worth \$ 149 per tonne fob .
-- [16 16 Paris 16] newsroom +331 4221 5432
1: 349303 Saudi_Arabia
2: 22989 Paris

```
3: 9317 European_Union
4: 11867 Germany
5: 11867 Germany
6: 349303 Saudi_Arabia
7: 9239 Europe
8: 11867 Germany
9: 1300085 Écu
10: 349303 Saudi_Arabia
11: 9239 Europe
12: 349303 Saudi_Arabia
13: 9317 European_Union
14: 11867 Germany
15: 1300085 Écu
16: 22989 Paris
1: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.51337165,
final_local_scores=0.4980703, pem=0.0, similarity=18.953125]
2: 22989 Paris
    [pem_pos=0, final_scores=0.51705724,
final_local_scores=0.52478886, pem=-0.115301244,
similarity=20.377485]
3: 9317 European_Union
    [pem_pos=0, final_scores=0.39005715,
final_local_scores=0.38737214, pem=-0.005814796,
similarity=15.681292]
4: 11867 Germany
    [pem_pos=1, final_scores=0.6758311,
final_local_scores=0.66329914, pem=-1.0465823,
similarity=29.575361]
5: 11867 Germany
    [pem_pos=0, final_scores=0.71812373,
final_local_scores=0.7094974, pem=-0.15529114,
similarity=26.107029]
6: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.44506878,
final_local_scores=0.42345983, pem=-0.25454172,
similarity=18.110836]
7: 9239 Europe
    [pem_pos=0, final_scores=0.4853093,
final_local_scores=0.45756882, pem=-0.25557294,
similarity=19.134325]
8: 11867 Germany
    [pem_pos=1, final_scores=0.7000314,
final_local_scores=0.6897343, pem=-1.0465823,
similarity=30.364237]
9: 1300085 Écu
    [pem_pos=0, final_scores=0.5822008,
final_local_scores=0.71170354, pem=0.0, similarity=25.328365]
10: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.43723935,
final_local_scores=0.4149074, pem=-0.25454172,
similarity=17.855616]
```

```
11: 9239 Europe
    [pem_pos=0, final_scores=0.4774372,
final_local_scores=0.44896972, pem=-0.25557294,
similarity=18.87771]
12: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.45714587,
final_local_scores=0.43665224, pem=0.0, similarity=17.120287]
13: 9317 European_Union
    [pem_pos=0, final_scores=0.49338335,
final_local_scores=0.5002404, pem=-0.005814796,
similarity=19.049505]
14: 11867 Germany
    [pem_pos=1, final_scores=0.6796204,
final_local_scores=0.6674383, pem=-1.0465823,
similarity=29.698883]
15: 1300085 Écu
    [pem_pos=0, final_scores=0.57068783,
final_local_scores=0.6991273, pem=0.0, similarity=24.953066]
16: 22989 Paris
    [pem_pos=0, final_scores=0.57693654,
final_local_scores=0.590198, pem=-0.115301244,
similarity=22.329424]
```

Sample: b'875_French'

[1 1 French 1] shipyard workers march against job cuts . [2 2 RENNES 2] 1996-08-29 About 3,500 naval shipyard workers marched in the centre of the northern port town of [3 3 Cherbourg 3] on Thursday to protest against defence restructuring , a union official said . The local police headquarters did not give a figure but said 1,800 workers at the [4 4 Cherbourg 4] yard had stopped work . A cutback plan could slim their numbers to 1,700 from 4,200 . Several hundred workers also marched in the western town of [5 5 Indre 5] where 500 or 1,600 jobs are at risk .

```
1: 5843419 France
2: 46132 Rennes
3: 105031 Cherbourg-Octeville
4: 105031 Cherbourg-Octeville
5: 90576 Indre
1: 5843419 France
    [pem_pos=0, final_scores=0.7076661,
final_local_scores=0.7264238, pem=-0.80223274,
similarity=30.130314]
2: 46132 Rennes
    [pem_pos=0, final_scores=0.30713612,
final_local_scores=0.27976742, pem=-0.33276722,
similarity=14.248171]
3: 105031 Cherbourg-Octeville
    [pem_pos=0, final_scores=0.34958392,
final_local_scores=0.3238446, pem=-0.14523119,
similarity=14.543672]
4: 105031 Cherbourg-Octeville
```

```
[pem_pos=0, final_scores=0.42828184,
final_local_scores=0.4098102, pem=-0.14523119,
similarity=17.109056]
5: 90576 Indre
[pem_pos=0, final_scores=0.2902698,
final_local_scores=0.33707616, pem=0.0, similarity=14.148738]
```

Sample: b'876_PRESS'

PRESS DIGEST - [1 1 Pakistan 1] - August 29 . Following are some of the main stories in Thursday 's [2 2 Pakistani 2] newspapers : [3 3 DAWN 3] - The government has decided to transfer the entire distribution network of electricity to foreign managements to curtail losses of billions of rupees . - The government has suffered a loss of 11 billion rupees due to tax holidays at industrial estates in Hub and Gadoon . - High Court officials have unearthed police-run human cages at [4 4 Tando Allahyar 4] near [5 5 Hyderabad 5] . Some 27 people were rescued from the private jail set up by the police . - Opposition leader [6 6 Nawaz Sharif 6] renewed a pledge to oust the Pakistan People 's Party government headed by Prime Minister [7 7 Benazir Bhutto 7] . [8 8 BUSINESS READER 8] - Gas prices may go up by five percent to increase the rate of return of [9 9 Sui Southern Gas 9] and Sui Northern Gas companies . - [10 10 Japan 10] is importing 80 percent of cotton yarn from [11 11 Pakistan 11] every year . - The government has blamed sugar technologists for not supporting a long-term programme of research and development to increase production of sugarcane . [12 12 FINANCIAL POST 12] - Armed robbers pillaged 70 barrels of crude oil from a well near [13 13 Gujar Khan 13] on Wednesday . - [14 14 Pakistan 14] will pay an additional bill of \$ 244 million as private power projects with capacity of 3,225 megawatt go on-line by 1998/99 . [15 15 THE NATION 15] - The government is facing extreme difficulties in meeting its revenue collections targets for 1996/97 . - Mohib Textile Mills has defaulted to nearly 23 development finance institutions , foreign and local banks , leasing companies and modarabas ([16 16 Islamic 16] mutual funds) . - Investment Minister [17 17 Asif Ali Zardari 17] expressed keenness for a close working relationship with [18 18 Japanese 18] companies so that investment from [19 19 Japan 19] can multiply.q - [20 20 Karachi Stock Exchange 20] index falls by 7.84 points . THE NEWS - The prime minister 's special economic assistant Shahid Hasan Khan said privatisation of thermal power plants , power generation from private plants and management contracts of Area Electricity Boards would help achieve 6.5 percent GDP growth . - [21 21 Pakistan 21] 's [22 22 Muslim Commercial Bank 22] , Vital Information System , and Duff and Phelps of the [23 23 U.S. 23] are likely to announce a strategic alliance with [24 24 Bangladesh 24] 's only credit rating company -- Credit Rating and Information Systems Ltd -- next month . - The [25 25 Sindh High Court 25] issued an ad-

interim order restraining the Privatisation Commission from handing over Javedan Cement to Dadabhoy Investment (pvt) Ltd until it can consider a legal challenge mounted by unions to the deal . THE MUSLIM - [26 26 Pakistan 26] and [27 27 Iran 27] have agreed to expand and strengthen political , trade and economic relations . -- [28 28 Islamabad 28] newsroom

9251-274757

- 1: 23235 Pakistan
- 2: 23235 Pakistan
- 3: 929985 Dawn_(newspaper)
- 4: 4175901 Tando_Allahyar
- 5: 288090 Hyderabad,_Sindh
- 6: 374012 Nawaz_Shariif
- 7: 193737 Benazir_Bhutto
- 8: 3536877 Business_Recorder
- 9: 7409873 Sui_Southern_Gas_Company
- 10: 15573 Japan
- 11: 23235 Pakistan
- 12: 1305365 Financial_Post
- 13: 2782484 Gujar_Khan
- 14: 23235 Pakistan
- 15: 39902584 The_Nation
- 16: 6037917 Islam
- 17: 993655 Asif_Ali_Zardari
- 18: 15573 Japan
- 19: 15573 Japan
- 20: 1466645 Karachi_Stock_Exchange
- 21: 23235 Pakistan
- 22: 4665846 MCB_Bank_Limited
- 23: 3434750 United_States
- 24: 3454 Bangladesh
- 25: 11042310 Sindh_High_Court
- 26: 23235 Pakistan
- 27: 14653 Iran
- 28: 51153 Islamabad
- 1: 23235 Pakistan
 - [pem_pos=0, final_scores=0.6316863, final_local_scores=0.588194, pem=-0.21944217, similarity=22.835957]
- 2: 23235 Pakistan
 - [pem_pos=0, final_scores=0.52471244, final_local_scores=0.47134143, pem=-0.49826378, similarity=20.865118]
- 3: 929985 Dawn_(newspaper)
 - [pem_pos=2, final_scores=0.48834914, final_local_scores=0.5154647, pem=-1.8698962, similarity=29.640993]
- 4: 4175901 Tando_Allahyar
 - [pem_pos=0, final_scores=0.40761226, final_local_scores=0.49109274, pem=0.0, similarity=18.7449]
- 5: 288090 Hyderabad,_Sindh
 - [pem_pos=2, final_scores=0.25747454,

```
final_local_scores=0.24134159, pem=-2.0006502,  
similarity=22.171675]  
6: 374012 Nawaz_Shari  
    [pem_pos=0, final_scores=0.5681043,  
final_local_scores=0.5911299, pem=0.0, similarity=21.730207]  
7: 193737 Benazir_Bhutto  
    [pem_pos=0, final_scores=0.41270292,  
final_local_scores=0.42220914, pem=0.0, similarity=16.689276]  
8: 3536877 Business_Recorder  
    [pem_pos=0, final_scores=0.40233046,  
final_local_scores=0.52323955, pem=0.0, similarity=19.704226]  
9: 7409873 Sui_Southern_Gas_Company  
    [pem_pos=0, final_scores=0.36790812,  
final_local_scores=0.46651065, pem=0.0, similarity=18.011322]  
10: 15573 Japan  
    [pem_pos=0, final_scores=0.64690727,  
final_local_scores=0.70202184, pem=-0.13549899,  
similarity=25.776308]  
11: 23235 Pakistan  
    [pem_pos=0, final_scores=0.611002,  
final_local_scores=0.5655996, pem=-0.21944217,  
similarity=22.161694]  
12: 1305365 Financial_Post  
    [pem_pos=0, final_scores=0.34277356,  
final_local_scores=0.45425278, pem=0.0, similarity=17.645521]  
13: 2782484 Gujar_Khan  
    [pem_pos=0, final_scores=0.39977598,  
final_local_scores=0.44520783, pem=0.0, similarity=17.375603]  
14: 23235 Pakistan  
    [pem_pos=0, final_scores=0.65954,  
final_local_scores=0.61862004, pem=-0.21944217,  
similarity=23.743929]  
15: 39902584 The_Nation  
    [pem_pos=0, final_scores=0.36337268,  
final_local_scores=0.44239014, pem=-0.015505094,  
similarity=17.375835]  
16: 6037917 Islam  
    [pem_pos=0, final_scores=0.3477931,  
final_local_scores=0.38156766, pem=-0.21141872,  
similarity=16.62618]  
17: 993655 Asif_Ali_Zardari  
    [pem_pos=0, final_scores=0.53065974,  
final_local_scores=0.5581213, pem=0.0, similarity=20.745167]  
18: 15573 Japan  
    [pem_pos=0, final_scores=0.8144892,  
final_local_scores=0.8850797, pem=-0.99076414,  
similarity=35.890186]  
19: 15573 Japan  
    [pem_pos=0, final_scores=0.5881699,  
final_local_scores=0.6378602, pem=-0.13549899,  
similarity=23.861597]  
20: 1466645 Karachi_Stock_Exchange
```

```
[pem_pos=0, final_scores=0.53094304,
final_local_scores=0.6162893, pem=0.0, similarity=22.481016]
21: 23235 Pakistan
    [pem_pos=0, final_scores=0.71517056,
final_local_scores=0.67938805, pem=-0.21944217,
similarity=25.557367]
22: 4665846 MCB_Bank_Limited
    [pem_pos=0, final_scores=0.5159556,
final_local_scores=0.5760221, pem=0.0, similarity=21.279362]
23: 3434750 United_States
    [pem_pos=0, final_scores=0.45885283,
final_local_scores=0.47006708, pem=-0.14075686,
similarity=18.882912]
24: 3454 Bangladesh
    [pem_pos=0, final_scores=0.5096358,
final_local_scores=0.48885894, pem=-0.28971618,
similarity=20.253761]
25: 11042310 Sindh_High_Court
    [pem_pos=0, final_scores=0.355206,
final_local_scores=0.41769552, pem=0.0, similarity=16.55458]
26: 23235 Pakistan
    [pem_pos=0, final_scores=0.5857901,
final_local_scores=0.5380595, pem=-0.21944217,
similarity=21.339842]
27: 14653 Iran
    [pem_pos=0, final_scores=0.600887,
final_local_scores=0.682696, pem=-0.13856095,
similarity=25.21624]
28: 51153 Islamabad
    [pem_pos=0, final_scores=0.3551795,
final_local_scores=0.31006113, pem=0.0, similarity=13.342555]
```

Sample: b'877_Salang'

[1 1 Salang 1] tunnel reopened linking [2 2 Kabul 2] with north . [3 3 SALANG 3] TUNNEL , [4 4 Afghanistan 4] 1996-08-29 The [5 5 Salang 5] tunnel linking [6 6 Kabul 6] with northern [7 7 Afghanistan 7] was formally reopened to traffic on Thursday under an agreement between the government and an opposition militia , witnesses said . They said dozens of trucks began moving through the tunnel from both directions after the road reopened . The [8 8 Salang 8] tunnel , the main supply route for [9 9 Soviet 9] troops when they were occupying [10 10 Afghanistan 10] in the 1980s , had been closed since 1994 when northern militia leader General [1 1 11 Abdul Rashid Dostum 11] rebelled against the [12 12 Kabul 12] government . Witnesses said wrecked tanks and vehicles littered both sides of the heavily-mined road . Mines had been removed from the road itself , but experts of the [13 13 Halo Trust 13] mine clearance agency said it would take a week to clear the roadsides . [14 14 Afghan 14] Deputy Prime Minister Qotbuddin Hilal officiated at the reopening ceremony , which was delayed by several hours while the two

sides argued about a mutual release of prisoners .

1: 18037690 Salang,_Afghanistan
2: 16826 Kabul
3: 18037690 Salang,_Afghanistan
4: 737 Afghanistan
5: 18037690 Salang,_Afghanistan
6: 16826 Kabul
7: 737 Afghanistan
8: 18037690 Salang,_Afghanistan
9: 26779 Soviet_Union
10: 737 Afghanistan
11: 2376 Abdul_Rashid_Dostum
12: 16826 Kabul
13: 385441 HALO_Trust
14: 737 Afghanistan

1: 18037690 Salang,_Afghanistan
[pem_pos=2, final_scores=0.58586776,
final_local_scores=0.68620926, pem=-1.867658,
similarity=34.72418]

2: 16826 Kabul
[pem_pos=0, final_scores=0.3564269,
final_local_scores=0.31636068, pem=0.0, similarity=13.530547]

3: 18037690 Salang,_Afghanistan
[pem_pos=2, final_scores=0.55520076,
final_local_scores=0.65271026, pem=-1.867658,
similarity=33.724503]

4: 737 Afghanistan
[pem_pos=0, final_scores=0.54012704,
final_local_scores=0.5091517, pem=-0.22286008,
similarity=20.495764]

5: 18037690 Salang,_Afghanistan
[pem_pos=2, final_scores=0.7612516,
final_local_scores=0.87778944, pem=-1.867658,
similarity=40.441315]

6: 16826 Kabul
[pem_pos=0, final_scores=0.35612202,
final_local_scores=0.3160277, pem=0.0, similarity=13.52061]

7: 737 Afghanistan
[pem_pos=0, final_scores=0.65854496,
final_local_scores=0.6385053, pem=-0.22286008,
similarity=24.355932]

8: 18037690 Salang,_Afghanistan
[pem_pos=2, final_scores=0.81339604,
final_local_scores=0.9347493, pem=-1.867658,
similarity=42.14111]

9: 26779 Soviet_Union
[pem_pos=0, final_scores=0.26395726,
final_local_scores=0.3180913, pem=-0.097171314,
similarity=14.110624]

10: 737 Afghanistan
[pem_pos=0, final_scores=0.56407714,
final_local_scores=0.53531355, pem=-0.22286008,

```
similarity=21.276485]
11: 2376 Abdul_Rashid_Dostum
    [pem_pos=0, final_scores=0.56407344,
final_local_scores=0.6299048, pem=0.0, similarity=22.88733]
12: 16826 Kabul
    [pem_pos=0, final_scores=0.37134522,
final_local_scores=0.3326567, pem=0.0, similarity=14.016852]
13: 385441 HALO_Trust
    [pem_pos=0, final_scores=0.43295163,
final_local_scores=0.5244155, pem=0.0, similarity=19.739317]
14: 737 Afghanistan
    [pem_pos=0, final_scores=0.49948066,
final_local_scores=0.46475166, pem=-0.869566,
similarity=22.687664]
```

Sample: b'878_Two'

Two [1 1 Indians 1] to die for killing 23 bus passengers . [2 2 NEW DELHI 2] 1996-08-29 [3 3 India 3] 's Supreme Court on Thursday sentenced two men to death after finding them guilty of killing 23 bus passengers , including children . It said the two , after robbing the passengers , burnt them alive by sprinkling the bus with petrol and setting it on fire in the southern state of [4 4 Andhra Pradesh 4] in 1993 . " We have no doubt that this is one of the rarest of the rare cases , not merely due to the number of innocent human beings roasted alive by the appellants , but the inhuman manner in which they plotted the scheme and executed it , " Justice K.T. Thomas said in the verdict by a panel of three judges .

```
1: 14533 India
2: 51585 New_Delhi
3: 14533 India
4: 2377 Andhra_Pradesh
1: 14533 India
    [pem_pos=2, final_scores=0.4294424,
final_local_scores=0.3659894, pem=-1.8953385,
similarity=25.318712]
2: 51585 New_Delhi
    [pem_pos=0, final_scores=0.375915,
final_local_scores=0.36343855, pem=0.0, similarity=14.935444]
3: 14533 India
    [pem_pos=0, final_scores=0.688473,
final_local_scores=0.648941, pem=-0.13622136,
similarity=24.1962]
4: 2377 Andhra_Pradesh
    [pem_pos=0, final_scores=0.38959783,
final_local_scores=0.36376184, pem=0.0, similarity=14.945091]
```

Sample: b'879_Elephant'

Elephant tramples woman to death in [1 1 Nepal 1] . [2 2 KAT HMANDU 2] 1996-08-29 A rampaging elephant dragged a sleeping 72-year-old woman from her bed and trampled her to death in the third such killing in two months , [3 3 Nepal 3] police

said on Thursday . The elephant crashed into Hari Maya Poudels house in [4 4 Madhumalla 4] village earlier this week while she was asleep , they said . The beast dragged the woman 30 feet (nine metres) away from her bed and trampled her to death , a police official told [5 5 Reuters 5] in the [6 6 Himalayan 6] kingdoms capital [7 7 Kathmandu 7] . In the past two months elephants have killed three people in remote areas of east and central [8 8 Nepal 8] . Elephants are protected under [9 9 Nepali 9] law , which provides for jail sentences of up to 15 years for convicted elephant killers .

1: 171166 Nepal
2: 17168 Kathmandu
3: 171166 Nepal
4: 20715492 Madhumalla
5: 18998750 Reuters
6: 4418894 Himalaya_Kingdom
7: 17168 Kathmandu
8: 171166 Nepal
9: 171166 Nepal
1: 171166 Nepal
[pem_pos=0, final_scores=0.65443164,
final_local_scores=0.6058492, pem=-0.22887385,
similarity=23.414114]
2: 17168 Kathmandu
[pem_pos=0, final_scores=0.428707,
final_local_scores=0.40103942, pem=0.0, similarity=16.05753]
3: 171166 Nepal
[pem_pos=0, final_scores=0.59698457,
final_local_scores=0.543097, pem=-0.22887385,
similarity=21.541462]
4: 20715492 Madhumalla
[pem_pos=0, final_scores=0.30537975,
final_local_scores=0.45711464, pem=0.0, similarity=17.730925]
5: 18998750 Reuters
[pem_pos=0, final_scores=0.34439015,
final_local_scores=0.45063072, pem=0.0, similarity=17.537432]
6: 4418894 Himalaya_Kingdom
[pem_pos=2, final_scores=0.097294465,
final_local_scores=0.24392965, pem=-1.5999681,
similarity=20.069935]
7: 17168 Kathmandu
[pem_pos=0, final_scores=0.42587745,
final_local_scores=0.39794856, pem=0.0, similarity=15.965292]
8: 171166 Nepal
[pem_pos=0, final_scores=0.62027365,
final_local_scores=0.5685368, pem=-0.22887385,
similarity=22.300636]
9: 171166 Nepal
[pem_pos=1, final_scores=0.54866964,
final_local_scores=0.49032027, pem=-1.4646683,
similarity=26.686941]

Sample: b'880_Sri'
[1 1 Sri Lankan 1] rebels overrun police post , kill 24 . [2
2 COLOMBO 2] 1996-08-29 [3 3 Tamil Tiger 3] rebels overran an
isolated police post in [4 4 Sri Lanka 4] 's northeast early
on Thursday killing 24 policemen , defence officials said . A
large group of [5 5 Liberation Tigers of Tamil Eelam 5] ([6 6
LTTE 6]) rebels stormed the Kudapokuna police post , just
north of Welikanda , 200 km (125 miles) from [7 7 Colombo 7]
, before dawn , they said . " The entire post was overrun , "
said a defence official . It was not immediately clear if
there were any casualties among the rebels , who are fighting
for independence for minority [8 8 Tamils 8] in the Indian
Ocean island 's north and east . It is the second time in
three days that the rebels attacked police . Suspected [9 9 Ta
mil 9] Tigers on Tuesday hurled hand grenades at a police
vehicle in a crowded market in the army-controlled northern
town of [10 10 Vavuniya 10] , killing at least two
policemen . More than a dozen people , including several
police who were working undercover , were wounded in the
attack . [11 11 Vavuniya 11] is just south of the northern
mainland area controlled by the [12 12 LTTE 12] . The
government says more than 50,000 people have died in the
ethnic war , now in its 14th year .

1: 26750 Sri_Lanka
2: 56636 Colombo
3: 18606 Liberation_Tigers_of_Tamil_Eelam
4: 26750 Sri_Lanka
5: 18606 Liberation_Tigers_of_Tamil_Eelam
6: 18606 Liberation_Tigers_of_Tamil_Eelam
7: 56636 Colombo
8: 300548 Tamil_people
9: 300548 Tamil_people
10: 4126071 Vavuniya
11: 4126071 Vavuniya
12: 18606 Liberation_Tigers_of_Tamil_Eelam
1: 26750 Sri_Lanka
 [pem_pos=0, final_scores=0.4939646,
final_local_scores=0.48442924, pem=-0.27403015,
similarity=20.036266]
2: 56636 Colombo
 [pem_pos=0, final_scores=0.49478,
final_local_scores=0.48156232, pem=-0.19393882,
similarity=19.515163]
3: 18606 Liberation_Tigers_of_Tamil_Eelam
 [pem_pos=0, final_scores=0.4697253,
final_local_scores=0.43109316, pem=0.0, similarity=16.954393]
4: 26750 Sri_Lanka
 [pem_pos=0, final_scores=0.54873,
final_local_scores=0.5442521, pem=-0.3027737,
similarity=21.977812]
5: 18606 Liberation_Tigers_of_Tamil_Eelam
 [pem_pos=0, final_scores=0.54680055,

```
final_local_scores=0.5152862, pem=0.0, similarity=19.46688]
6: 18606 Liberation_Tigers_of_Tamil_Eelam
    [pem_pos=0, final_scores=0.4636628,
final_local_scores=0.42447084, pem=0.0, similarity=16.75677]
7: 56636 Colombo
    [pem_pos=0, final_scores=0.4106068,
final_local_scores=0.38961583, pem=-0.19393882,
similarity=16.771296]
8: 300548 Tamil_people
    [pem_pos=0, final_scores=0.4023173,
final_local_scores=0.39827102, pem=-0.005057526,
similarity=16.002419]
9: 300548 Tamil_people
    [pem_pos=1, final_scores=0.13231252,
final_local_scores=0.10897839, pem=-1.7722582,
similarity=16.979662]
10: 4126071 Vavuniya
    [pem_pos=0, final_scores=0.39725614,
final_local_scores=0.41238445, pem=0.0, similarity=16.396088]
11: 4126071 Vavuniya
    [pem_pos=0, final_scores=0.37083977,
final_local_scores=0.3835286, pem=0.0, similarity=15.534971]
12: 18606 Liberation_Tigers_of_Tamil_Eelam
    [pem_pos=0, final_scores=0.5081291,
final_local_scores=0.4730435, pem=0.0, similarity=18.206274]
```

Sample: b'881_Vicorp'

Vicorp Restaurants names Sabourin CFO . [1 1 DENVER 1]
1996-08-29 Vicorp Restaurants Inc said it has named Richard
Sabourin as executive vice president and chief financial
officer . The company said Sabourin is the former president
and chief executive at Bestop Inc of [2 2 Boulder 2] , 1
Colo . It said Craig Held has also joined the company as
executive vice president and chief marketing officer . -- New
York Newsdesk 212 859 1610

1: 8522 Denver

2: 94341 Boulder,_Colorado

1: 5399 Colorado

[winner: 114241 Colo,_Iowa, pem_pos=0,
final_scores=-0.20423117, final_local_scores=-0.1348536,
pem=-0.522792, similarity=2.9084365|
gt: 5399 Colorado, pem_pos=13, final_scores=-0.51320434,
final_local_scores=-0.61852986, pem=-4.919981,
similarity=12.38717]

[winner: 4721653 Zyphe_language, pem_pos=12,
final_scores=-1.6050379, final_local_scores=-1.5283818,
pem=-4.4188404, similarity=-17.48991|
RECALL MISS: 5399 Colorado]

1: 8522 Denver

[pem_pos=0, final_scores=0.53458965,
final_local_scores=0.5093062, pem=-0.23925911,
similarity=20.589554]

```
2: 94341 Boulder,_Colorado
    [pem_pos=0, final_scores=0.5348248,
final_local_scores=0.5095631, pem=-0.49585652,
similarity=21.992638]

Sample: b'882_SoCal'
SoCal Edison sees 2 power lines back today . [1 1 NEW YORK 1]
1996-08-29 Southern California Edison Co said it expected two
220 kilovolt ( KV ) power lines in southern [2 2 California 2]
to resume service later today after being shut late Wednesday
because of a wildfire raging north of [3 3 Los Angeles 3] .
They are expected to be placed in service later today ,
said company spokesman Steve Conroy , adding repair crews
have been removing smoke and other fire-related residues ,
which had settled on the two lines . The shutdown of the 220
KV lines reduced by 500 megawatts ( MW ) the amount of power
which the area received from SoCal Edison 's 1,200 MW [4 4 Si
erra 4] hydroelectric facility , he said . Conroy noted two
500 KV and another 220 KV line running from the [5 5 Sierra 5]
plant to [6 6 Los Angeles 6] remained in operation , and
continued to carry some of the production from [7 7 Sierra 7]
to the region . On Monday , the two 500 KV transmission
cables were taken out of service , also for cleaning , for
about a day . Containment of the fire has been difficult
because of the hot , arid , windy weather in the region ,
Conroy said . " The fires keep moving back because of the
winds , " he said , forcing the utility to shut those
transmission lines for a second time this week . Local
authorities charged a teenager for starting the blaze . In
four days , the fire destroyed 20,000 acres of forest land .
-- R Leong , New York Power Desk +1 212 859 1622
1: 645042 New_York_City
2: 5407 California
3: 18110 Los_Angeles
4: 50413 Sierra_Nevada_(U.S.)
5: 50413 Sierra_Nevada_(U.S.)
6: 18110 Los_Angeles
7: 50413 Sierra_Nevada_(U.S.)
1: 645042 New_York_City
    [pem_pos=1, final_scores=0.42634833,
final_local_scores=0.4286816, pem=-1.7906687,
similarity=26.620361]
2: 5407 California
    [pem_pos=0, final_scores=0.6052874,
final_local_scores=0.56504875, pem=-0.08662026,
similarity=21.42295]
3: 18110 Los_Angeles
    [pem_pos=0, final_scores=0.53637487,
final_local_scores=0.5312856, pem=-0.100652285,
similarity=20.491697]
4: 50413 Sierra_Nevada_(U.S.)
    [pem_pos=3, final_scores=0.5025521,
```

```
final_local_scores=0.49719232, pem=-2.3475947,  
similarity=31.69351]  
5: 50413 Sierra_Nevada_(U.S.)  
    [pem_pos=3, final_scores=0.5145317,  
final_local_scores=0.51027817, pem=-2.3475947,  
similarity=32.084015]  
6: 18110 Los_Angeles  
    [pem_pos=0, final_scores=0.5338738,  
final_local_scores=0.5285535, pem=-0.100652285,  
similarity=20.410166]  
7: 50413 Sierra_Nevada_(U.S.)  
    [pem_pos=3, final_scores=0.51848596,  
final_local_scores=0.5145977, pem=-2.3475947,  
similarity=32.21292]
```

Sample: b'883_Dreyfus'

Dreyfus Strategic Munis monthly \$ 0.056 / shr . [1 1 NEW YORK
1] 1996-08-29 Monthly Latest Prior Amount \$ 0.056 \$ 0.056 Pay
Sept 27 Record Sept 13 NOTE : Full name of company is Dreyfus
Strategic Municipal Inc .

1: 8210131 New_York

1: 8210131 New_York

```
    [pem_pos=0, final_scores=0.10903408,  
final_local_scores=0.39169037, pem=-0.2889045,  
similarity=17.349642]
```

Sample: b'884_Sierra'

Sierra Semiconductor jumps on exit plan . Martin Wolk [1 1 SE
ATTLE 1] 1996-08-29 Sierra Semiconductor Corp jumped 23
percent Thursday on the expectation the company would emerge
as a smaller but more profitable operation after its planned
exit from the computer modem business . The [2 2 San Jose 2] ,
[3 3 Calif. 3] , company was up 2-1/8 at 11-3/8 after its
announcement Wednesday that it planned to pull out of the
highly competitive modem-chip business and focus instead on
the fast-growing market for computer networking equipment . "Cer-
tainly the company will be a much smaller company now , but it will be a more profitable business , " said analyst Elias Moosa of Roberston Stephens & Co . But analysts noted that Sierra still has much painful work ahead of it , including cutting as many as 150 jobs from its workforce , which currently has 500 people , and building up the business of its PMC-Sierra unit , which makes routing devices and chipsets for high-speed computer networks . The company has announced plans to take a charge against earnings of \$ 50 million to \$ 80 million to write down the value of assets and inventories and cover severance payments . Scott Randall of Soundview Financial Group said the company likely would have difficulty selling its modem-chip business . "Once you announce your intention to exit a business , it becomes a complete buyer 's market , " he said . And he said that while the company is focusing on the fastest-growing part of its

business , the market for networking chips has begun to attract the attention of much-larger players such as International Business Machines Corp . " As the market develops the question is , are they able to make that transition to be a much larger company ? " Randall said . Other analysts were more bullish , even though the company is expected to shrink to slightly more than half its current size in sales . " It 's a positive strategic move , " said Miles Kan of [4 4 Hambrecht & Quist 4] . " The modem business is a low-margin , commodity business , " he said . The company 's [5 5 PMC-Sierra 5] unit generated \$ 33 million of the company 's \$ 117 million in sales in the first half of the year , compared with \$ 45 million in sales of modem chips , Kan said . But the PMC unit is far more profitable , he said . Sierra 's stock has fallen from a high of nearly \$ 25 this year as the computer chip sector has been battered by falling prices and concern about slowing demand . -- [6 6 Seattle 6] bureau 206-386-4848

1: 11388236 Seattle
2: 53446 San_Jose,_California
3: 5407 California
4: 17476047 Hambrecht_&_Quist
5: 790369 PMC-Sierra
6: 11388236 Seattle
1: 11388236 Seattle
[pem_pos=0, final_scores=0.5751518,
final_local_scores=0.5520333, pem=-0.1054736,
similarity=21.13707]
2: 53446 San_Jose,_California
[pem_pos=0, final_scores=0.39144188,
final_local_scores=0.35663873, pem=-0.3689591,
similarity=16.738981]
3: 5407 California
[pem_pos=0, final_scores=0.63922364,
final_local_scores=0.64886534, pem=0.0, similarity=23.45315]
4: 17476047 Hambrecht_&_Quist
[pem_pos=0, final_scores=0.36132503,
final_local_scores=0.4477842, pem=0.0, similarity=17.452486]
5: 790369 PMC-Sierra
[pem_pos=0, final_scores=0.4089008,
final_local_scores=0.53149205, pem=0.0, similarity=19.950497]
6: 11388236 Seattle
[pem_pos=0, final_scores=0.57640725,
final_local_scores=0.5534046, pem=-0.1054736,
similarity=21.177994]

Sample: b'885_Housecall'
Housecall shares sink after profit warning . [1 1 NEW YORK 1]
1996-08-29 Shares in home healthcare services company
Housecall Medical Resources Inc fell more than 50 percent on
Thursday after the company said it expected a net loss for
the fiscal fourth quarter . [2 2 Morgan Stanley 2] said it

downgraded the stock to underperform from outperform . Housecall was off 7-3/8 to 7-1/8 in morning trading . The Atlanta-based company went public in April at \$ 16 a share . [3 3 Wall Street 3] had expected the company to earn \$ 0.17 a share in its fourth quarter , ended June 30 , according to First Call . Housecall said fourth quarter earnings and revenues were expected to fall short of expectations . It said its non-Medicare infusion therapy , hospice and nursing services businesses failed to meet budgeted revenues . It also cited a limitation on [4 4 Medicare 4] reimbursement for some services provided during the quarter .

```
1: 645042 New_York_City
2: 74384 Morgan_Stanley
3: 21560 New_York_Stock_Exchange
4: 58353 Medicare_(United_States)
1: 645042 New_York_City
    [pem_pos=1, final_scores=0.38287932,
final_local_scores=0.42575443, pem=-1.7906687,
similarity=26.533009]
2: 74384 Morgan_Stanley
    [pem_pos=0, final_scores=0.46183413,
final_local_scores=0.48464417, pem=0.0, similarity=18.552462]
3: 21560 New_York_Stock_Exchange
    [pem_pos=7, final_scores=0.48410076,
final_local_scores=0.49936044, pem=-2.2964394,
similarity=31.48002]
4: 58353 Medicare_(United_States)
    [pem_pos=0, final_scores=0.34229475,
final_local_scores=0.45011395, pem=0.0, similarity=17.52201]
```

Sample: b'886_First'

First Alliance net income slips . [1 1 IRVINE 1] , [2 2 Cali f. 2] 1996-08-29 First Alliance Corporation and Subsidiaries Quarter ended June 30 Six Months ended June 30 1996 1995 1996 1995 Unaudited Total revenue 17,024,000 18,174,000 31,834,000 24,137,000 Total expense 7,718,000 6,828,000 14,668,000 13,091,000 Net Income 9,167,000 11,175,000 16,909,000 10,880,000 Net Income Per Share 0.86 1.05 1.59 1.02 Weighted average number of shares outstanding 10,650,407 10,650,407 10,650,407 10,650,407 Pro Forma : Historical income before income tax provision 9,306,000 11,346,000 17,166,000 11,046,000 Pro forma income tax provision 3,820,000 4,658,000 7,047,000 4,534,000 Pro forma net income 5,486,000 6,688,000 10,119,000 6,512,000 Pro forma net income per share 0.37 0.45 0.68 0.44 Pro forma weighted average number of shares outstanding 14,775,000 14,775,000 14,775,000 14,775,000

```
1: 5201333 Irvine,_California
2: 5407 California
1: 5201333 Irvine,_California
    [pem_pos=0, final_scores=0.41330987,
final_local_scores=0.41902602, pem=-0.6463034,
similarity=20.108984]
```

```
2: 5407 California
    [pem_pos=0, final_scores=0.5455497,
final_local_scores=0.56347793, pem=0.0, similarity=20.905018]
```

Sample: b'887_Oasis'

[1 1 Oasis 1] singer heads for [2 2 U.S. 2] after illness . [3 3 LONDON 3] 1996-08-29 [4 4 Liam Gallagher 4] , singer of [5 5 Britain 5] 's top rock group [6 6 Oasis 6] , flew out on Thursday to join the band three days after the start of its [7 7 U.S. 7] tour . [8 8 Gallagher 8] made his usual obscene gestures and swore at journalists as he prepared to fly from [9 9 London 9] 's [10 10 Heathrow 10] airport to [11 11 Chicago 11] . " I hate you f... ing lot , yet you 're always asking me too many things . I 'm not a supermodel you know , " he said . On Monday , just 15 minutes before his flight was due to depart , [12 12 Liam 12] decided not to travel with the rest of the group , which includes his brother [13 13 Noel 13] . [14 14 Liam 14] caught a taxi back to [15 15 London 15] saying he had " problems at home " . He was believed to be suffering from laryngitis and said he had to go house-hunting with actress girlfriend [16 16 Patsy Kensit 16] . When they first heard that [17 17 Liam 17] had not flown out with the band at the start of the tour , many [18 18 U.S. 18] fans asked for refunds on their concert tickets . The group began the [19 19 U.S. 19] tour , which is scheduled to last until September 18 , with a concert in [20 20 Chicago 20] on Tuesday at which [21 21 Noel Gallagher 21] filled in for his brother as lead singer .

```
1: 161515 Oasis_(band)
2: 3434750 United_States
3: 17867 London
4: 178326 Liam_Gallagher
5: 31717 United_Kingdom
6: 161515 Oasis_(band)
7: 3434750 United_States
8: 178326 Liam_Gallagher
9: 17867 London
10: 13595 London_Heathrow_Airport
11: 6886 Chicago
12: 178326 Liam_Gallagher
13: 194961 Noel_Gallagher
14: 178326 Liam_Gallagher
15: 17867 London
16: 882447 Patsy_Kensit
17: 178326 Liam_Gallagher
18: 3434750 United_States
19: 3434750 United_States
20: 6886 Chicago
21: 194961 Noel_Gallagher
1: 161515 Oasis_(band)
    [pem_pos=0, final_scores=0.4449914,
final_local_scores=0.41727, pem=-0.2557975,
```

```
similarity=17.93295]
2: 3434750 United_States
    [pem_pos=0, final_scores=0.41799027,
final_local_scores=0.44459575, pem=-0.14075686,
similarity=18.122793]
3: 17867 London
    [pem_pos=0, final_scores=0.6602893,
final_local_scores=0.68182313, pem=-0.09086737,
similarity=24.930826]
4: 178326 Liam_Gallagher
    [pem_pos=0, final_scores=0.48634535,
final_local_scores=0.45201725, pem=0.0, similarity=17.57881]
5: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5318584,
final_local_scores=0.5516949, pem=-0.61114657,
similarity=23.8769]
6: 161515 Oasis_(band)
    [pem_pos=0, final_scores=0.43962723,
final_local_scores=0.41141045, pem=-0.2557975,
similarity=17.75809]
7: 3434750 United_States
    [pem_pos=0, final_scores=0.4430141,
final_local_scores=0.47193044, pem=-0.14075686,
similarity=18.938515]
8: 178326 Liam_Gallagher
    [pem_pos=1, final_scores=0.8861477,
final_local_scores=0.8955589, pem=0.0, similarity=30.814978]
9: 17867 London
    [pem_pos=0, final_scores=0.6567626,
final_local_scores=0.67797077, pem=-0.09086737,
similarity=24.815863]
10: 13595 London_Heathrow_Airport
    [pem_pos=0, final_scores=0.3501209,
final_local_scores=0.4021933, pem=0.0, similarity=16.091963]
11: 6886 Chicago
    [pem_pos=0, final_scores=0.6049634,
final_local_scores=0.6381044, pem=-0.20658566,
similarity=24.255466]
12: 178326 Liam_Gallagher
    [pem_pos=0, final_scores=0.7173255,
final_local_scores=0.7089072, pem=0.0, similarity=25.244917]
13: 194961 Noel_Gallagher
    [pem_pos=0, final_scores=0.77577955,
final_local_scores=0.7790741, pem=0.0, similarity=27.338837]
14: 178326 Liam_Gallagher
    [pem_pos=0, final_scores=0.7255016,
final_local_scores=0.7178383, pem=0.0, similarity=25.511438]
15: 17867 London
    [pem_pos=0, final_scores=0.57998526,
final_local_scores=0.5941032, pem=-0.09086737,
similarity=22.313087]
16: 882447 Patsy_Kensit
```

```
[pem_pos=0, final_scores=0.5222022,
final_local_scores=0.5562374, pem=0.0, similarity=20.688946]
17: 178326 Liam_Gallagher
    [pem_pos=0, final_scores=0.68477374,
final_local_scores=0.67334926, pem=0.0, similarity=24.183798]
18: 3434750 United_States
    [pem_pos=0, final_scores=0.4470356,
final_local_scores=0.4763233, pem=-0.14075686,
similarity=19.069609]
19: 3434750 United_States
    [pem_pos=0, final_scores=0.45907962,
final_local_scores=0.4894796, pem=-0.14075686,
similarity=19.462217]
20: 6886 Chicago
    [pem_pos=0, final_scores=0.52995694,
final_local_scores=0.55617124, pem=-0.20658566,
similarity=21.810417]
21: 194961 Noel_Gallagher
    [pem_pos=0, final_scores=0.47051436,
final_local_scores=0.44054228, pem=0.0, similarity=17.236374]
```

Sample: b'888_Slough'

[1 1 Slough Estates 1] helps lift property sector . [2 2 LON DON 2] 1996-08-29 A strong set of interim results and an upbeat outlook from Slough Estates Plc helped to boost the property sector on Thursday . Shares in Slough , which earlier announced a 14 percent rise in first-half pretax profit to 37.4 million stg , climbed nearly six percent , or 14p to 250 pence at 1009 [3 3 GMT 3] , while [4 4 British Land 4] added 12.1 / 2p to 468p , [5 5 Land Securities 5] rose 5.1 / 2p to 691p and [6 6 Hammerson 6] was 8p higher at 390 . Traders said positive comment from investment banks [7 7 Merrill Lynch 7] and SBC Warburg also fueled the gains . One dealer said positive stances from [8 8 Merrill Lynch 8] and SBC Warburg were the key factors behind the gains . A spokesman for [9 9 Merrill Lynch 9] said the bank was preparing to issue a note on the sector . " We have been very positive (on property) , " he said , adding : " On a technical basis it is our most favoured sector . " SBC Warburg issued an update on the property sector on Thursday , saying that most of the predictions it made at the start of the year were being realised . " In the property market it is a case of so far , so good , " a member of SBC Warburg 's property team said . SBC Warburg said it is maintaining its forecast for five percent growth in rental incomes during 1996 , but it has shaved its forecast for capital growth to five percent from six . The spokesman said SBC Warburg has also put an " add " recommendation on [10 10 Slough Estates 10] ' shares , but added that this " is a general move , not because of the results . " Slough 's chairman Sir [11 11 Nigel Mobbs 11] added to the bullish mood in the sector , saying in a statement that " with the prospect of a period of steady

economic growth and low inflation ahead , there is good reason to believe that the property sector should continue its improvement . " -- Jonathan Birt , London Newsroom +44 171 542 7717

1: 7028227 Segro
2: 17867 London
3: 12701 Greenwich_Mean_Time
4: 1771497 British_Land
5: 2049638 Land_Securities
6: 2928064 Hammerson
7: 33731299 Merrill_Lynch
8: 33731299 Merrill_Lynch
9: 33731299 Merrill_Lynch
10: 7028227 Segro
11: 2587583 Nigel_Mobbs

1: 7028227 Segro
[pem_pos=0, final_scores=0.42310876,
final_local_scores=0.45394087, pem=0.0, similarity=17.636213]
2: 17867 London
[pem_pos=0, final_scores=0.6215375,
final_local_scores=0.6961379, pem=-0.09086737,
similarity=25.358007]
3: 12701 Greenwich_Mean_Time
[pem_pos=0, final_scores=0.57447916,
final_local_scores=0.7291536, pem=0.0, similarity=25.84911]
4: 1771497 British_Land
[pem_pos=0, final_scores=0.5624847,
final_local_scores=0.5906748, pem=0.0, similarity=21.716627]
5: 2049638 Land_Securities
[pem_pos=0, final_scores=0.51834553,
final_local_scores=0.5564033, pem=0.0, similarity=20.693897]
6: 2928064 Hammerson
[pem_pos=0, final_scores=0.69517976,
final_local_scores=0.7378023, pem=0.0, similarity=26.107206]
7: 33731299 Merrill_Lynch
[pem_pos=0, final_scores=0.51516134,
final_local_scores=0.5278134, pem=0.0, similarity=19.840717]
8: 33731299 Merrill_Lynch
[pem_pos=0, final_scores=0.51666504,
final_local_scores=0.52945596, pem=0.0, similarity=19.889734]
9: 33731299 Merrill_Lynch
[pem_pos=0, final_scores=0.51567155,
final_local_scores=0.52837074, pem=0.0, similarity=19.85735]
10: 7028227 Segro
[pem_pos=0, final_scores=0.5007414,
final_local_scores=0.53874284, pem=0.0, similarity=20.166874]
11: 2587583 Nigel_Mobbs
[pem_pos=0, final_scores=0.23502634,
final_local_scores=0.36593544, pem=0.0, similarity=15.009956]

Sample: b'889_Canada'
[1 1 Canada 1] 's international travel account gap shrinks .

[2 2 OTTAWA 2] 1996-08-29 Higher spending by foreign visitors and less [3 3 Canadian 3] tourist spending abroad cut the deficit in [4 4 Canada 4]'s international travel account by 26.5 percent in the second quarter , [5 5 Statistics Canada 5] said on Thursday . The deficit fell to a seasonally adjusted C\$ 715 million in the second quarter from C\$ 973 million in the first , as foreigners spent a record C\$ 3.00 billion seasonally adjusted while [6 6 Canadians 6] reduced their spending abroad by 5.1 percent to C\$ 3.72 billion . --
Reuters Ottawa Burea (613) 235-6745

1: 5042916 Canada
2: 22219 Ottawa
3: 5042916 Canada
4: 5042916 Canada
5: 186760 Statistics_Canada
6: 5042916 Canada
1: 5042916 Canada
[pem_pos=0, final_scores=0.8040785,
final_local_scores=0.75964004, pem=-0.11662793,
similarity=27.393127]
2: 22219 Ottawa
[pem_pos=0, final_scores=0.47185475,
final_local_scores=0.43827313, pem=-0.34965137,
similarity=19.070116]
3: 5042916 Canada
[pem_pos=0, final_scores=0.5976263,
final_local_scores=0.5341224, pem=-0.41303828,
similarity=22.275156]
4: 5042916 Canada
[pem_pos=0, final_scores=0.6427404,
final_local_scores=0.58340275, pem=-0.11662793,
similarity=22.133856]
5: 186760 Statistics_Canada
[pem_pos=0, final_scores=0.48872238,
final_local_scores=0.54460067, pem=0.0, similarity=20.341682]
6: 5042916 Canada
[pem_pos=0, final_scores=0.5328011,
final_local_scores=0.46331072, pem=-0.8178484,
similarity=22.363415]

Sample: b'890_Jordanian'
[1 1 Jordanian 1] PM Kabariti meets [2 2 Arafat 2] in [3 3 We
st Bank 3] . [4 4 RAMALLAH 4] , [5 5 West Bank 5] 1996-08-29 [
6 6 Jordanian 6] Prime Minister Abdul-Karim al-Kabariti began
talks with [7 7 Palestinian 7] President [8 8 Yasser Arafat 8]
in the [9 9 West Bank 9] on Thursday on the stalled [10 10 Mi
ddle East 10] peace process , officials said . Kabariti flew
by helicopter to Palestinian-ruled [11 11 Ramallah 11] and
after a brief arrival ceremony went into talks with [12 12 Ar
afat 12] . The prime minister 's visit , his first trip
outside the country since [13 13 Jordan 13] was shaken by
food riots earlier this month , came against the backdrop of

a [14 14 Palestinian 14] general strike in the [15 15 West Bank 15] and [16 16 Gaza 16] . [17 17 Arafat 17] called the four-hour strike , which ended at noon (0900 [18 18 GMT 18]) to protest against [19 19 Israeli 19] policy on settlements and [20 20 Jerusalem 20] . [21 21 Jordan 21] 's official state news agency [22 22 Petra 22] said Kabariti would hold discussions " on the latest developments in the peace process and bilateral cooperation " .

1: 7515964 Jordan
2: 34289 Yasser_Arafat
3: 96665 Palestinian_territories
4: 44132 Ramallah
5: 96665 Palestinian_territories
6: 7515964 Jordan
7: 241405 State_of_Palestine
8: 34289 Yasser_Arafat
9: 96665 Palestinian_territories
10: 19323 Middle_East
11: 44132 Ramallah
12: 34289 Yasser_Arafat
13: 7515964 Jordan
14: 241405 State_of_Palestine
15: 96665 Palestinian_territories
16: 12047 Gaza_Strip
17: 34289 Yasser_Arafat
18: 12701 Greenwich_Mean_Time
19: 9282173 Israel
20: 16043 Jerusalem
21: 7515964 Jordan
22: 6226592 Jordan_News_Agency
1: 7515964 Jordan
 [pem_pos=0, final_scores=0.51711637,
final_local_scores=0.47413492, pem=-0.18608934,
similarity=19.250828]
2: 34289 Yasser_Arafat
 [pem_pos=0, final_scores=0.655086,
final_local_scores=0.63513863, pem=0.0, similarity=23.043516]
3: 96665 Palestinian_territories
 [pem_pos=1, final_scores=0.49385422,
final_local_scores=0.4495629, pem=-1.929408,
similarity=27.997988]
4: 44132 Ramallah
 [pem_pos=0, final_scores=0.45161217,
final_local_scores=0.41373825, pem=0.0, similarity=16.436487]
5: 96665 Palestinian_territories
 [pem_pos=1, final_scores=0.35605037,
final_local_scores=0.2990331, pem=-1.929408,
similarity=23.505877]
6: 7515964 Jordan
 [pem_pos=0, final_scores=0.5515819,
final_local_scores=0.5117833, pem=-0.18608934,
similarity=20.37433]

7: 241405 State_of_Palestine
[pem_pos=9, final_scores=0.7315224,
final_local_scores=0.73584247, pem=-2.8508017,
similarity=41.551823]
8: 34289 Yasser_Arafat
[pem_pos=0, final_scores=0.6391699,
final_local_scores=0.6135167, pem=0.0, similarity=22.398275]
9: 96665 Palestinian_territories
[pem_pos=1, final_scores=0.39892662,
final_local_scores=0.3458689, pem=-1.929408,
similarity=24.903551]
10: 19323 Middle_East
[pem_pos=0, final_scores=0.50251955,
final_local_scores=0.5051992, pem=0.0, similarity=19.165865]
11: 44132 Ramallah
[pem_pos=0, final_scores=0.44035167,
final_local_scores=0.40143788, pem=0.0, similarity=16.06942]
12: 34289 Yasser_Arafat
[pem_pos=0, final_scores=0.58665174,
final_local_scores=0.56038463, pem=0.0, similarity=20.812708]
13: 7515964 Jordan
[pem_pos=0, final_scores=0.5607535,
final_local_scores=0.52180195, pem=-0.53647494,
similarity=22.578758]
14: 241405 State_of_Palestine
[pem_pos=9, final_scores=0.646684,
final_local_scores=0.6431695, pem=-2.8508017,
similarity=38.786278]
15: 96665 Palestinian_territories
[pem_pos=1, final_scores=0.4280507,
final_local_scores=0.3776825, pem=-1.929408,
similarity=25.852932]
16: 12047 Gaza_Strip
[pem_pos=1, final_scores=0.3196311,
final_local_scores=0.27421868, pem=-1.374985,
similarity=19.750328]
17: 34289 Yasser_Arafat
[pem_pos=0, final_scores=0.60636264,
final_local_scores=0.5819158, pem=0.0, similarity=21.45524]
18: 12701 Greenwich_Mean_Time
[pem_pos=0, final_scores=0.6282285,
final_local_scores=0.7396461, pem=0.0, similarity=26.162228]
19: 9282173 Israel
[pem_pos=0, final_scores=0.5986031,
final_local_scores=0.567604, pem=-0.34790626,
similarity=22.920118]
20: 16043 Jerusalem
[pem_pos=0, final_scores=0.4883837,
final_local_scores=0.4797092, pem=-0.13239029,
similarity=19.12515]
21: 7515964 Jordan
[pem_pos=0, final_scores=0.559344,

```
final_local_scores=0.5202622, pem=-0.53647494,  
similarity=22.532808]  
22: 6226592 Jordan_News_Agency  
    [pem_pos=15, final_scores=0.61562043,  
final_local_scores=0.7559964, pem=-4.3721313,  
similarity=50.426483]
```

Sample: b'891_Iran'

[1 1 Iran 1] says five spy networks destroyed , 41 held . [2
2 TEHRAN 2] 1996-08-29 [3 3 Iranian 3] security forces have
broken up five espionage rings in northwestern [4 4 Iran 4]
and arrested 41 people on charges of spying for unnamed
countries , a daily newspaper said on Thursday . Jomhuri
Eslami quoted the [5 5 West Azerbaijan 5] province security
chief as saying those held confessed to gathering
confidential information , photographing strategic sites ,
doing propaganda against state officials and " spreading [6 6
pan-Turkism 6]" . It was not clear if they were the same
five spy rings , allegedly led by Turkish diplomats , that [7
7 Iran 7] said in April it had broken up in the same area ,
which borders Turkey . The April arrests were announced
shortly after a row in which [8 8 Tehran 8] asked [9 9 Ankara
9] to withdraw four Turkish diplomats accused of spying , and
Turkey expelled four [10 10 Iranian 10] diplomats for their
alleged links to killings of [11 11 Iranian 11] exiles . Ties
between the two neighbours , strained also over a military
accord between Turkey and [12 12 Israel 12] which drew strong
[13 13 Iranian 13] objections , have improved since [14 14 Is
lamist 14] [15 15 Necmettin Erbakan 15] took over as Turkish
prime minister in June . The daily [16 16 Iran 16] on
Thursday quoted Intelligence Minister Ali Fallahiyan as
saying agents arrested 137 people for allegedly spying for [17
17 Iraq 17] , the [18 18 United States 18] and other unnamed
countries in the [19 19 Iranian 19] year which ended on March
19 .

1: 14653 Iran
2: 57654 Tehran
3: 14653 Iran
4: 14653 Iran
5: 530357 West_Azerbaijan_Province
6: 860964 Pan-Turkism
7: 14653 Iran
8: 57654 Tehran
9: 802 Ankara
10: 14653 Iran
11: 14653 Iran
12: 9282173 Israel
13: 14653 Iran
14: 6037917 Islam
15: 1328699 Necmettin_Erbakan
16: 14653 Iran
17: 7515928 Iraq

```
18: 3434750 United_States
19: 14653 Iran
1: 14653 Iran
    [pem_pos=0, final_scores=0.6353565,
final_local_scores=0.59382856, pem=-0.13856095,
similarity=22.564259]
2: 57654 Tehran
    [pem_pos=0, final_scores=0.45900828,
final_local_scores=0.4590972, pem=0.0, similarity=17.790089]
3: 14653 Iran
    [pem_pos=0, final_scores=0.6382064,
final_local_scores=0.59694165, pem=-0.57066226,
similarity=25.006992]
4: 14653 Iran
    [pem_pos=0, final_scores=0.6046931,
final_local_scores=0.56033343, pem=-0.13856095,
similarity=21.564697]
5: 530357 West_Azerbaijan_Province
    [pem_pos=0, final_scores=0.337456,
final_local_scores=0.39914078, pem=0.0, similarity=16.00087]
6: 860964 Pan-Turkism
    [pem_pos=0, final_scores=0.4368543,
final_local_scores=0.5260713, pem=0.0, similarity=19.78873]
7: 14653 Iran
    [pem_pos=0, final_scores=0.69873995,
final_local_scores=0.6630654, pem=-0.13856095,
similarity=24.630423]
8: 57654 Tehran
    [pem_pos=0, final_scores=0.24045514,
final_local_scores=0.22036114, pem=0.0, similarity=10.665729]
9: 802 Ankara
    [pem_pos=0, final_scores=0.32742512,
final_local_scores=0.35243604, pem=0.0, similarity=14.607107]
10: 14653 Iran
    [pem_pos=0, final_scores=0.5533344,
final_local_scores=0.5042319, pem=-0.57066226,
similarity=22.240349]
11: 14653 Iran
    [pem_pos=0, final_scores=0.59862405,
final_local_scores=0.55370396, pem=-0.57066226,
similarity=23.716694]
12: 9282173 Israel
    [pem_pos=0, final_scores=0.52924365,
final_local_scores=0.56364363, pem=-0.21279564,
similarity=22.06718]
13: 14653 Iran
    [pem_pos=0, final_scores=0.6335282,
final_local_scores=0.5918314, pem=-0.57066226,
similarity=24.854492]
14: 6037917 Islam
    [pem_pos=1, final_scores=0.32228357,
final_local_scores=0.35129377, pem=-1.3455799,
```

```
similarity=21.890491]
15: 1328699 Necmettin_Erbakan
    [pem_pos=0, final_scores=0.5030598,
final_local_scores=0.6324897, pem=0.0, similarity=22.964466]
16: 14653 Iran
    [pem_pos=0, final_scores=0.61453813,
final_local_scores=0.5710877, pem=-0.13856095,
similarity=21.885626]
17: 7515928 Iraq
    [pem_pos=0, final_scores=0.5874838,
final_local_scores=0.59331506, pem=-0.25379938,
similarity=23.17562]
18: 3434750 United_States
    [pem_pos=0, final_scores=0.47996116,
final_local_scores=0.52749103, pem=-0.073457584,
similarity=20.230572]
19: 14653 Iran
    [pem_pos=0, final_scores=0.51588947,
final_local_scores=0.46332896, pem=-0.57066226,
similarity=21.019722]
```

Sample: b'892_Mideast'

Mideast Gulf oil outlook - [1 1 India 1] holds the key . [2 2 DUBAI 2] 1996-08-29 [3 3 India 3] will continue to hold the key to the middle distillates product market in the Middle East Gulf in the short term , traders in the region said on Thursday . They said premiums on high quality jet kerosene have widened to around \$ 1 , and are likely to remain strong in the near term . " On the jet kerosene side we must be cautious about quality . Some are commanding a very good premium of 95 cents to one dollar . It 's been 99 cents to [4 4 Korea 4] . This will stay at this sort of price premium for a while , " one said . " But for normal jet the weakened demand is quite noticeable , the premium is around 45-50 cents . The differential is not normally as wide as this , " he added . Another put the premium for jet kerosene at between 65 and 75 cents . One trader said despite the fact that some kerosene was in storage at [5 5 Dubai 5] ports , demand looked to exceed supply in the near term . Jet kerosene was assessed at \$ 27.40 - \$ 27.70 a barrel fob Gulf on Thursday , up from \$ 27.22 last week . Dealers expected premiums to stick around current levels for the next two or three weeks , before they get a boost in the second half of September from demand for October cargoes . Gas oil , assessed at \$ 24.00 - \$ 24.20 a barrel fob Gulf , was little changed on Thursday from last week 's \$ 24.10 - \$ 24.24 . " On gas oil , in the near term demand and supply are balanced to a bit short , there are some enquiries into east [6 6 Africa 6] and short covering in [7 7 India 7] , " said one trader . " The [8 8 Indians 8] have awarded three cargoes , but the question is whether they will come out for more , " he added . [9 9 India 9] has acquired 120,000 tonnes of

diesel in three cargoes , bound for the west coast , in its October tender . Traders said the award could be [10 10 India 10] 's lowest in recent years . " After [11 11 IOC 11] 's very small purchase of 120,000 tonnes , I 'm still suspicious that they will buy more in the second half of October , " another said . But traders see the market remaining tight in the short-term , with some surplus arising closer to October . " I see an overhang of gas oil further out , " one said . " Gas oil will remain tight in the short term up to mid-September with the premium of around \$ 1 for 0.5 percent (sulphur material) . In the second half of September and October we see the premium coming off to 70-75 cents . " Traders said there was not much demand for one percent sulphur material gas oil , with the premium at 40 to 45 cents .

```
1: 14533 India
2: 211583 Dubai
3: 14533 India
4: 16749 Korea
5: 211583 Dubai
6: 5334607 Africa
7: 14533 India
8: 14533 India
9: 14533 India
10: 14533 India
11: 7971634 Indian_Oil_Corporation
1: 14533 India
    [pem_pos=0, final_scores=0.66872185,
final_local_scores=0.61615, pem=-0.13622136,
similarity=23.217651]
2: 211583 Dubai
    [pem_pos=0, final_scores=0.47307664,
final_local_scores=0.48827618, pem=-0.08179128,
similarity=19.105642]
3: 14533 India
    [pem_pos=0, final_scores=0.58609164,
final_local_scores=0.52588904, pem=-0.13622136,
similarity=20.524084]
4: 16749 Korea
    [pem_pos=0, final_scores=0.3116567,
final_local_scores=0.3175796, pem=-0.27902022,
similarity=15.084276]
5: 211583 Dubai
    [pem_pos=0, final_scores=0.35957354,
final_local_scores=0.36429125, pem=-0.08179128,
similarity=15.405685]
6: 5334607 Africa
    [pem_pos=0, final_scores=0.6363595,
final_local_scores=0.65152365, pem=-0.13197383,
similarity=24.250174]
7: 14533 India
    [pem_pos=0, final_scores=0.62689644,
```

```
final_local_scores=0.5704621, pem=-0.13622136,
similarity=21.854233]
8: 14533 India
    [pem_pos=2, final_scores=0.5409511,
final_local_scores=0.47657984, pem=-1.8953385,
similarity=28.618952]
9: 14533 India
    [pem_pos=0, final_scores=0.5832467,
final_local_scores=0.5227814, pem=-0.13622136,
similarity=20.431345]
10: 14533 India
    [pem_pos=0, final_scores=0.64886755,
final_local_scores=0.5944622, pem=-0.13622136,
similarity=22.570444]
11: 7971634 Indian_Oil_Corporation
    [pem_pos=2, final_scores=0.66562694,
final_local_scores=0.7914807, pem=-1.36746,
similarity=35.14554]
```

Sample: b'893_Egypt'

[1 1 Egypt 1] police catch ancient manuscript thieves . [2 2 CAIRO 2] 1996-08-29 [3 3 Egyptian 3] police have arrested eight people who were trying to sell an ancient copy of the [4 4 Old Testament 4] , the official [5 5 al-Akhbar 5] newspaper said on Thursday . The daily said the men had wanted to sell the undated manuscript to a [6 6 Jewish 6] group for five million pounds (\$ 1.5 million) . Instead an undercover police officer pretended to be interested in buying it and arrested them . The newspaper did not give any details about the manuscript but said it had been relinquished to the [7 7 Islamic Museum 7] in [8 8 Cairo 8] .

```
1: 8087628 Egypt
2: 6293 Cairo
3: 8087628 Egypt
4: 22326 Old_Testament
5: 7614001 Al_Akhbar_(Lebanon)
6: 25955086 Jews
7: 19623898 Islamic_Museum,_Jerusalem
8: 6293 Cairo
1: 8087628 Egypt
    [pem_pos=0, final_scores=0.57302696,
final_local_scores=0.5420587, pem=-0.27480945,
similarity=21.760283]
2: 6293 Cairo
    [pem_pos=0, final_scores=0.60898656,
final_local_scores=0.5884946, pem=-0.15211283,
similarity=22.478779]
3: 8087628 Egypt
    [pem_pos=0, final_scores=0.5331545,
final_local_scores=0.4985041, pem=-1.052751,
similarity=24.691093]
4: 22326 Old_Testament
```

```
[pem_pos=0, final_scores=0.357845,
final_local_scores=0.40688366, pem=0.0, similarity=16.231934]
5: 7614001 Al_Akhbar_(Lebanon)
    [pem_pos=0, final_scores=0.55921465,
final_local_scores=0.6708311, pem=0.0, similarity=24.10865]
6: 25955086 Jews
    [pem_pos=1, final_scores=0.3878271,
final_local_scores=0.38560027, pem=-0.80476934,
similarity=19.973255]
7: 19623898 Islamic_Museum,_Jerusalem
    [pem_pos=0, final_scores=0.3506444,
final_local_scores=0.4262644, pem=0.0, similarity=16.810293]
8: 6293 Cairo
    [pem_pos=0, final_scores=0.4236712,
final_local_scores=0.38606572, pem=-0.15211283,
similarity=16.437899]
```

Sample: b'894_New'

```
New [1 1 U.S. 1] ambassador arrives in [2 2 Saudi Arabia 2] .
[3 3 DUBAI 3] 1996-08-29 [4 4 Washington 4] 's new
ambassador to [5 5 Saudi Arabia 5] , Wyche [6 6 Fowler 6] ,
arrived in the kingdom to take up his post , the [7 7 U.S. 7]
embassy in [8 8 Riyadh 8] said on Thursday . [9 9 Fowler 9] ,
a lawyer and former senator , arrived late on Wednesday , the
embassy said in a statement . President [10 10 Bill Clinton 10]
earlier this month invoked special powers to appoint [11 11
Fowler 11] during the congressional recess because the Senate
delayed confirming his nomination . [12 12 Fowler 12] 's
predecessor Raymond Mabus returned to the [13 13 United States
13] in May .
```

```
1: 3434750 United_States
2: 349303 Saudi_Arabia
3: 211583 Dubai
4: 13015878 Washington_(state)
5: 349303 Saudi_Arabia
6: 909551 Wyche_Fowler
7: 3434750 United_States
8: 51554 Riyadh
9: 909551 Wyche_Fowler
10: 3356 Bill_Clinton
11: 909551 Wyche_Fowler
12: 909551 Wyche_Fowler
13: 3434750 United_States
1: 3434750 United_States
    [pem_pos=0, final_scores=0.54422945,
final_local_scores=0.5279593, pem=-0.14075686,
similarity=20.610529]
2: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.4693964,
final_local_scores=0.46552783, pem=-0.25454172,
similarity=19.36623]
3: 211583 Dubai
```

```
[pem_pos=0, final_scores=0.44564432,
final_local_scores=0.46770608, pem=-0.08179128,
similarity=18.491789]
4: 13015878 Washington_(state)
    [pem_pos=1, final_scores=0.3268556,
final_local_scores=0.35146192, pem=-1.9417106,
similarity=25.137362]
5: 349303 Saudi_Arabia
    [pem_pos=0, final_scores=0.4348951,
final_local_scores=0.4278404, pem=-0.25454172,
similarity=18.24156]
6: 909551 Wyche_Fowler
    [pem_pos=19, final_scores=0.51637983,
final_local_scores=0.5438753, pem=-3.838158,
similarity=41.192535]
7: 3434750 United_States
    [pem_pos=0, final_scores=0.5627367,
final_local_scores=0.54817563, pem=-0.14075686,
similarity=21.213823]
8: 51554 Riyadh
    [pem_pos=0, final_scores=0.32564038,
final_local_scores=0.32457393, pem=0.0, similarity=13.775646]
9: 909551 Wyche_Fowler
    [pem_pos=0, final_scores=0.6940754,
final_local_scores=0.73798066, pem=0.0, similarity=26.112528]
10: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.50654525,
final_local_scores=0.5288555, pem=0.0, similarity=19.871817]
11: 909551 Wyche_Fowler
    [pem_pos=0, final_scores=0.62424976,
final_local_scores=0.66170675, pem=0.0, similarity=23.836363]
12: 909551 Wyche_Fowler
    [pem_pos=0, final_scores=0.7151446,
final_local_scores=0.7609956, pem=0.0, similarity=26.799341]
13: 3434750 United_States
    [pem_pos=0, final_scores=0.51165247,
final_local_scores=0.49237388, pem=-0.073457584,
similarity=19.182606]
```

Sample: b'895_Jordanian'

```
[1 1 Jordanian 1] PM Kabariti leaves for [2 2 West Bank 2] .
[3 3 AMMAN 3] 1996-08-29 [4 4 Jordanian 4] Prime Minister
Abdul-Karim al-Kabariti left [5 5 Amman 5] on Thursday for
the [6 6 West Bank 6] town of [7 7 Ramallah 7] to hold talks
with [8 8 Palestinian 8] President [9 9 Yasser Arafat 9] on
the stalled [10 10 Middle East 10] peace process , officials
said . The official state news agency [11 11 Petra 11] said
Kabariti would hold discussions " on the latest developments
in the peace process and bilateral cooperation " . The visit
was Kabariti 's first trip outside the country since [12 12 Jo
rdan 12] was shaken by food riots earlier this month .
1: 7515964 Jordan
```

```
2: 96665 Palestinian_territories
3: 45007 Amman
4: 7515964 Jordan
5: 45007 Amman
6: 96665 Palestinian_territories
7: 44132 Ramallah
8: 241405 State_of_Palestine
9: 34289 Yasser_Arafat
10: 19323 Middle_East
11: 6226592 Jordan_News_Agency
12: 7515964 Jordan
1: 7515964 Jordan
    [pem_pos=0, final_scores=0.50354314,
final_local_scores=0.44437438, pem=-0.18608934,
similarity=18.362713]
2: 96665 Palestinian_territories
    [pem_pos=1, final_scores=0.520686,
final_local_scores=0.48186976, pem=-1.929408,
similarity=28.962088]
3: 45007 Amman
    [pem_pos=0, final_scores=0.53029555,
final_local_scores=0.5008632, pem=0.0, similarity=19.03647]
4: 7515964 Jordan
    [pem_pos=0, final_scores=0.5263576,
final_local_scores=0.46929574, pem=-0.18608934,
similarity=19.106417]
5: 45007 Amman
    [pem_pos=0, final_scores=0.41183352,
final_local_scores=0.3714614, pem=0.0, similarity=15.174862]
6: 96665 Palestinian_territories
    [pem_pos=1, final_scores=0.45623785,
final_local_scores=0.41146994, pem=-1.929408,
similarity=26.861217]
7: 44132 Ramallah
    [pem_pos=0, final_scores=0.46548402,
final_local_scores=0.43910772, pem=0.0, similarity=17.193563]
8: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.7448297,
final_local_scores=0.7652558, pem=-2.8508017,
similarity=42.429577]
9: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.57874185,
final_local_scores=0.59419346, pem=0.0, similarity=21.82163]
10: 19323 Middle_East
    [pem_pos=0, final_scores=0.52234876,
final_local_scores=0.5127653, pem=0.0, similarity=19.391651]
11: 6226592 Jordan_News_Agency
    [pem_pos=15, final_scores=0.72944653,
final_local_scores=0.874693, pem=-4.3721313,
similarity=53.968628]
12: 7515964 Jordan
    [pem_pos=0, final_scores=0.5934213,
```

```
final_local_scores=0.5425526, pem=-0.53647494,
similarity=23.197998]

Sample: b'896_Palestinians'
[1 1 Palestinians 1] end four-hour strike . [2 2 JERUSALEM 2]
1996-08-29 [3 3 Palestinians 3] reopened their shops on
Thursday at the end of a four-hour strike called by President
[4 4 Yasser Arafat 4] to protest against [5 5 Israel 5] 's
policy on [6 6 Jewish 6] settlements and [7 7 Jerusalem 7] ,
witnesses said . Shopkeepers in Arab East Jerusalem rolled up
their shutters some 10 minutes before the scheduled noon
( 0900 [8 8 GMT 8] ) end of the stoppage . [9 9 Palestinian 9]
leaders called the strike , the first in the [1 West Bank 1]
and [10 10 Gaza 10] since 1994 , a warning signal that the
peace process with [11 11 Israel 11] was in danger .
Witnesses said most shops were closed in towns and villages
in the areas , with the exception of [12 12 Hebron 12] , a
West Bank city still under [13 13 Israeli 13] occupation .
1: 241405 State_of_Palestine
2: 16043 Jerusalem
3: 241405 State_of_Palestine
4: 34289 Yasser_Arafat
5: 9282173 Israel
6: 25955086 Jews
7: 16043 Jerusalem
8: 12701 Greenwich_Mean_Time
9: 241405 State_of_Palestine
10: 12047 Gaza_Strip
11: 9282173 Israel
12: 38577 Hebron
13: 9282173 Israel
1: 241405 State_of_Palestine
    [pem_pos=5, final_scores=0.5325591,
final_local_scores=0.517054, pem=-1.7909594,
similarity=29.259151]
2: 16043 Jerusalem
    [pem_pos=0, final_scores=0.5255591,
final_local_scores=0.497037, pem=-0.13239029,
similarity=19.642246]
3: 241405 State_of_Palestine
    [pem_pos=5, final_scores=0.5074262,
final_local_scores=0.48960012, pem=-1.7909594,
similarity=28.439873]
4: 34289 Yasser_Arafat
    [pem_pos=0, final_scores=0.579763,
final_local_scores=0.5812695, pem=0.0, similarity=21.435955]
5: 9282173 Israel
    [pem_pos=0, final_scores=0.6166435,
final_local_scores=0.57376975, pem=-0.21279564,
similarity=22.369362]
6: 25955086 Jews
    [pem_pos=1, final_scores=0.3732429,
```

```
final_local_scores=0.36855996, pem=-0.80476934,
similarity=19.464737]
7: 16043 Jerusalem
    [pem_pos=0, final_scores=0.60810286,
final_local_scores=0.5872035, pem=-0.13239029,
similarity=22.332996]
8: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.63780683,
final_local_scores=0.75527483, pem=0.0, similarity=26.62862]
9: 241405 State_of_Palestine
    [pem_pos=9, final_scores=0.6928351,
final_local_scores=0.69213116, pem=-2.8508017,
similarity=40.24739]
10: 12047 Gaza_Strip
    [pem_pos=1, final_scores=0.2774111,
final_local_scores=0.23384768, pem=-1.374985,
similarity=18.545576]
11: 9282173 Israel
    [pem_pos=0, final_scores=0.6437109,
final_local_scores=0.60333675, pem=-0.21279564,
similarity=23.251701]
12: 38577 Hebron
    [pem_pos=0, final_scores=0.41680676,
final_local_scores=0.40453702, pem=-0.39021417,
similarity=18.283949]
13: 9282173 Israel
    [pem_pos=0, final_scores=0.5271003,
final_local_scores=0.47595745, pem=-0.34790626,
similarity=20.1852]
1: 96665 Palestinian_territories
    [pem_pos=1, final_scores=0.050476983,
final_local_scores=-0.021161646, pem=-1.929408,
similarity=13.950627]
```

Sample: b'897_U.S.'

[1 1 U.S. 1] wheat weekly export sales highlights - [2 2 USDA
2] . [3 3 WASHINGTON 3] 1996-08-29 [4 4 U.S. 4] wheat major
net sales activity in the week ended Aug 22 reported by
exporters for the following purchasing countries , in
tonnes : Net Sales : 1996/97 1997/98 [5 5 Egypt 5] 199,900
Nil [6 6 S. Korea 6] 149,100-A [7 7 Japan 7] 74,600 [8 8 China
8] 55,000-B [9 9 Indonesia 9] 55,000-B Unknown - 161,600 A-
includes 54,600 tonnes changed from unknown . B- reflects
55,000 tonnes changed from unknown . Primary Export
Destinations : [10 10 Egypt 10] , [11 11 Morocco 11] , [12 12
S. Korea 12] , [13 13 Yemen 13] , [14 14 Pakistan 14] , [15 15
Mexico 15] and [16 16 China 16] .

1: 3434750 United_States
2: 70896 United_States_Department_of_Agriculture
3: 108956 Washington,_D.C.
4: 3434750 United_States
5: 8087628 Egypt

```
6: 27019 South_Korea
7: 15573 Japan
8: 5405 China
9: 14579 Indonesia
10: 8087628 Egypt
11: 19291 Morocco
12: 27019 South_Korea
13: 350939 Yemen
14: 23235 Pakistan
15: 3966054 Mexico
16: 5405 China
1: 3434750 United_States
    [pem_pos=0, final_scores=0.537882,
final_local_scores=0.50400525, pem=-0.14075686,
similarity=19.895693]
2: 70896 United_States_Department_of_Agriculture
    [pem_pos=0, final_scores=0.4113028,
final_local_scores=0.49564767, pem=0.0, similarity=18.880829]
3: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.5292271,
final_local_scores=0.5518362, pem=-1.7496481,
similarity=30.072462]
4: 3434750 United_States
    [pem_pos=0, final_scores=0.53093606,
final_local_scores=0.49641794, pem=-0.14075686,
similarity=19.669271]
5: 8087628 Egypt
    [pem_pos=0, final_scores=0.53860694,
final_local_scores=0.52239656, pem=-0.27480945,
similarity=21.173525]
6: 27019 South_Korea
    [pem_pos=0, final_scores=0.46542144,
final_local_scores=0.4504823, pem=0.0, similarity=17.533003]
7: 15573 Japan
    [pem_pos=0, final_scores=0.6852067,
final_local_scores=0.6822604, pem=-0.13549899,
similarity=25.186588]
8: 5405 China
    [pem_pos=0, final_scores=0.531163,
final_local_scores=0.49594343, pem=-0.15669827,
similarity=19.741804]
9: 14579 Indonesia
    [pem_pos=0, final_scores=0.52791804,
final_local_scores=0.5143228, pem=-0.23151328,
similarity=20.697138]
10: 8087628 Egypt
    [pem_pos=0, final_scores=0.4903379,
final_local_scores=0.46966994, pem=-0.27480945,
similarity=19.600058]
11: 19291 Morocco
    [pem_pos=0, final_scores=0.5177137,
final_local_scores=0.5034564, pem=-0.29598162,
```

```
similarity=20.72345]
12: 27019 South_Korea
    [pem_pos=0, final_scores=0.4991873,
final_local_scores=0.48736638, pem=0.0, similarity=18.633698]
13: 350939 Yemen
    [pem_pos=0, final_scores=0.5027283,
final_local_scores=0.4910041, pem=-0.27215755,
similarity=20.22229]
14: 23235 Pakistan
    [pem_pos=0, final_scores=0.4934953,
final_local_scores=0.4754187, pem=-0.21944217,
similarity=19.470514]
15: 3966054 Mexico
    [pem_pos=0, final_scores=0.7663968,
final_local_scores=0.7824479, pem=-0.18258883,
similarity=28.432465]
16: 5405 China
    [pem_pos=0, final_scores=0.48184997,
final_local_scores=0.44207644, pem=-0.15669827,
similarity=18.134306]
```

Sample: b'898_Keane'

```
[1 1 Keane 1] wins contract from [2 2 ING 2] units . [3 3 BO
STON 3] 1996-08-29 Software services company [4 4 Keane Inc 4]
said it had won a year 2000 compliance contract from Life
Insurance Co of Georgia and Southland Life Insurance Co ,
both part of [5 5 The Netherlands 5] ' [6 6 ING Group 6] .
The company said in a statement late on Wednesday it would
conduct an enterprise assessment and strategic compliance
plan for preparing all the mainframe systems of Life of
Georgia and Southland Life to operate in the new century .
The project will be managed by [7 7 Keane 7] 's [8 8 Atlanta 8
] office . " The client 's goal is to complete its year 2000
conversion activities by the end of 1997 , " the statement
said . -- [9 9 New York 9] newsroom , ( 212 ) 859-1610
1: 7712661 Keane_(company)
2: 230965 ING_Group
3: 24437894 Boston
4: 7712661 Keane_(company)
5: 21148 Netherlands
6: 230965 ING_Group
7: 7712661 Keane_(company)
8: 3138 Atlanta
9: 645042 New_York_City
1: 7712661 Keane_(company)
    [pem_pos=3, final_scores=0.57365006,
final_local_scores=0.6223748, pem=-3.2059393,
similarity=40.097015]
2: 230965 ING_Group
    [pem_pos=0, final_scores=0.47880423,
final_local_scores=0.51330143, pem=0.0, similarity=19.407652]
3: 24437894 Boston
```

```
[pem_pos=0, final_scores=0.63107747,
final_local_scores=0.6701751, pem=-0.23164093,
similarity=25.348774]
4: 7712661 Keane_(company)
    [pem_pos=0, final_scores=0.593339,
final_local_scores=0.6438819, pem=0.0, similarity=23.304434]
5: 21148 Netherlands
    [pem_pos=0, final_scores=0.5605849,
final_local_scores=0.6246695, pem=-0.10774545,
similarity=23.317032]
6: 230965 ING_Group
    [pem_pos=0, final_scores=0.49130046,
final_local_scores=0.5269517, pem=0.0, similarity=19.815002]
7: 7712661 Keane_(company)
    [pem_pos=3, final_scores=0.48401123,
final_local_scores=0.52445793, pem=-3.2059393,
similarity=37.17498]
8: 3138 Atlanta
    [pem_pos=0, final_scores=0.38266367,
final_local_scores=0.40448958, pem=-0.23152597,
similarity=17.419561]
9: 645042 New_York_City
    [pem_pos=1, final_scores=0.4080984,
final_local_scores=0.42234325, pem=-1.7906687,
similarity=26.431213]
```

Sample: b'899_PRESS'

PRESS DIGEST - [1 1 Wall Street Journal 1] - Aug 29 . [2 2 NE
W YORK 2] 1996-08-29 The [3 3 National Basketball Association
3] has sued America Online Inc , alleging that the [4 4 United
States 4] ' No 1 on-line service is delivering real-time
information about league games without its permission , [5 5 T
he Wall Street Journal 5] reported on Thursday . The suit was
filed on Wednesday in federal court in [6 6 Manhattan 6] and
is another legal skirmish over what constitutes a " broadcast
" in the computer age . The suit contends [7 7 America Online
7] was misappropriating [8 8 NBA 8] property by providing a
site containing continually updated scores and statistics of [
9 9 NBA 9] games in progress . The newspaper also reported :
* Baxter International Inc has reached an agreement to
acquire [10 10 Austria 10] 's Immuno International AG in a
complex deal valued at \$ 715 million . * [11 11 Boeing Co 11]
secures \$ 5.5 billion in orders for new , larger 747s . *
President [12 12 Bill Clinton 12] is expected to propose a
tax break on home sales . * Philip Morris Cos Inc raises
dividend 20 percent . * [13 13 Salomon Brothers Inc 13]
analyst is bullish on International Business Machines Corp .
* Sierra Semiconductor Corp puts modem chipset line up for
sale and sets layoffs . * Red Lion Hotels Inc says it 's
holding merger talks with Doubletree Corp . * GTE Corp , Baby
Bells and their allies ready to launch challenge to
telecommunications reform law . * Economists see second-

quarter gross domestic product revised down 0.1 percentage point . * President [14 14 Clinton 14] proposes five-point plan to clean up toxic waste sites . * Securities and Exchange Commission acts to improve stock-trade prices for investors . * H&R Block Inc delays spinoff of its stake in CompuServe Corp . * Stock funds see cash pour in again in July . * The Federal Trade Commission and Justice Department issue new guidelines for formation of doctors ' networks . -- [15 15 New York 15] newsroom , (212) 859-1610

1: 173070 The_Wall_Street_Journal
2: 645042 New_York_City
3: 22093 National_Basketball_Association
4: 3434750 United_States
5: 173070 The_Wall_Street_Journal
6: 45470 Manhattan
7: 1397 AOL
8: 22093 National_Basketball_Association
9: 22093 National_Basketball_Association
10: 26964606 Austria
11: 18933266 Boeing
12: 3356 Bill_Clinton
13: 1043102 Salomon_Brothers
14: 3356 Bill_Clinton
15: 645042 New_York_City
1: 173070 The_Wall_Street_Journal
 [pem_pos=0, final_scores=0.3679301,
final_local_scores=0.36471212, pem=0.0, similarity=14.97345]
2: 645042 New_York_City
 [pem_pos=1, final_scores=0.43537194,
final_local_scores=0.39387357, pem=-1.7906687,
similarity=25.581621]
3: 22093 National_Basketball_Association
 [pem_pos=0, final_scores=0.34314126,
final_local_scores=0.3492778, pem=0.0, similarity=14.512859]
4: 3434750 United_States
 [pem_pos=0, final_scores=0.5236824,
final_local_scores=0.5177216, pem=-0.073457584,
similarity=19.939032]
5: 173070 The_Wall_Street_Journal
 [pem_pos=0, final_scores=0.5027742,
final_local_scores=0.5120087, pem=0.0, similarity=19.369074]
6: 45470 Manhattan
 [pem_pos=0, final_scores=0.3225817,
final_local_scores=0.30270597, pem=-0.13335879,
similarity=13.848289]
7: 1397 AOL
 [pem_pos=0, final_scores=0.47098047,
final_local_scores=0.5065165, pem=0.0, similarity=19.205175]
8: 22093 National_Basketball_Association
 [pem_pos=0, final_scores=0.36962605,
final_local_scores=0.3782084, pem=0.0, similarity=15.376205]
9: 22093 National_Basketball_Association

```
[pem_pos=0, final_scores=0.34395963,
final_local_scores=0.35017177, pem=0.0, similarity=14.539537]
10: 26964606 Austria
    [pem_pos=0, final_scores=0.6186994,
final_local_scores=0.6899586, pem=-0.25262418,
similarity=26.053263]
11: 18933266 Boeing
    [pem_pos=0, final_scores=0.36030704,
final_local_scores=0.40234298, pem=0.0, similarity=16.09643]
12: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.5676976,
final_local_scores=0.56950766, pem=0.0, similarity=21.084957]
13: 1043102 Salomon_Brothers
    [pem_pos=0, final_scores=0.31312966,
final_local_scores=0.37922305, pem=0.0, similarity=15.406485]
14: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.8290797,
final_local_scores=0.86812776, pem=0.0, similarity=29.996378]
15: 645042 New_York_City
    [pem_pos=1, final_scores=0.45780277,
final_local_scores=0.41837585, pem=-1.7906687,
similarity=26.312817]
```

Sample: b'900_Baker'

[1 1 Baker 1] made secret trip to [2 2 Syria 2] in March 1995 . [3 3 WASHINGTON 3] 1996-08-29 Former Secretary of State [4 4 James Baker 4] made a secret trip to [5 5 Syria 5] in March 1995 in an unsuccessful bid to break an impasse in negotiations between [6 6 Syria 6] and [7 7 Israel 7] , the [8 8 Washington Post 8] reported on Thursday . The paper said [9 9 Baker 9] declined to discuss the trip , but authorised an associate to confirm it took place and give an account of it . News of the secret trip came after [10 10 Baker 10] trashed the [11 11 Clinton 11] administration at the [12 12 Republican National Convention 12] two weeks ago for its efforts to nudge [13 13 Syria 13] into peace with [14 14 Israel 14] . [15 15 Baker 15] made the March 1995 trip on the explicit understanding that it remain a secret , but after his speech at the [16 16 GOP 16] convention , [17 17 Israel 17] 's outgoing ambassador [18 18 Itamar Rabinovich 18] told a reporter about it , the [19 19 Post 19] said . [20 20 Baker 20] was secretary of state in the [21 21 Republican 21] administration of President [22 22 George Bush 22] .

```
1: 213137 James_Baker
2: 7515849 Syria
3: 108956 Washington,_D.C.
4: 213137 James_Baker
5: 7515849 Syria
6: 7515849 Syria
7: 9282173 Israel
8: 102226 The_Washington_Post
9: 213137 James_Baker
```

```
10: 213137 James_Baker
11: 3356 Bill_Clinton
12: 509001 Republican_National_Convention
13: 7515849 Syria
14: 9282173 Israel
15: 213137 James_Baker
16: 32070 Republican_Party_(United_States)
17: 9282173 Israel
18: 17944843 Itamar_Rabinovich
19: 102226 The_Washington_Post
20: 213137 James_Baker
21: 32070 Republican_Party_(United_States)
22: 11955 George_H._W._Bush
1: 213137 James_Baker
    [pem_pos=0, final_scores=0.5848849,
final_local_scores=0.5744636, pem=0.0, similarity=21.232853]
2: 7515849 Syria
    [pem_pos=0, final_scores=0.497602,
final_local_scores=0.51970315, pem=-0.26162508,
similarity=21.02145]
3: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.557025,
final_local_scores=0.55872625, pem=-1.7496481,
similarity=30.278076]
4: 213137 James_Baker
    [pem_pos=0, final_scores=0.4963041,
final_local_scores=0.47770256, pem=0.0, similarity=18.34531]
5: 7515849 Syria
    [pem_pos=0, final_scores=0.49961638,
final_local_scores=0.5219036, pem=-0.26162508,
similarity=21.087114]
6: 7515849 Syria
    [pem_pos=0, final_scores=0.60827905,
final_local_scores=0.6406011, pem=-0.26162508,
similarity=24.629286]
7: 9282173 Israel
    [pem_pos=0, final_scores=0.5503801,
final_local_scores=0.5510903, pem=-0.21279564,
similarity=21.692564]
8: 102226 The_Washington_Post
    [pem_pos=0, final_scores=0.5552533,
final_local_scores=0.60417086, pem=0.0, similarity=22.119377]
9: 213137 James_Baker
    [pem_pos=0, final_scores=0.5879093,
final_local_scores=0.5777673, pem=0.0, similarity=21.331442]
10: 213137 James_Baker
    [pem_pos=0, final_scores=0.5561423,
final_local_scores=0.5430666, pem=0.0, similarity=20.295904]
11: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.48664957,
final_local_scores=0.46654278, pem=-1.920079,
similarity=28.45397]
```

```
12: 509001 Republican_National_Convention
    [pem_pos=0, final_scores=0.3105256,
final_local_scores=0.33363402, pem=-0.21607539,
similarity=15.221069]
13: 7515849 Syria
    [pem_pos=0, final_scores=0.6421613,
final_local_scores=0.6776123, pem=-0.26162508,
similarity=25.733774]
14: 9282173 Israel
    [pem_pos=0, final_scores=0.5807451,
final_local_scores=0.58425945, pem=-0.21279564,
similarity=22.682396]
15: 213137 James_Baker
    [pem_pos=0, final_scores=0.63255066,
final_local_scores=0.6265312, pem=0.0, similarity=22.786654]
16: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.37669998,
final_local_scores=0.39143527, pem=0.0, similarity=15.770922]
17: 9282173 Israel
    [pem_pos=0, final_scores=0.632876,
final_local_scores=0.64120454, pem=-0.21279564,
similarity=24.38175]
18: 17944843 Itamar_Rabinovich
    [pem_pos=0, final_scores=0.4370137,
final_local_scores=0.5320261, pem=0.0, similarity=19.966433]
19: 102226 The_Washington_Post
    [pem_pos=6, final_scores=0.24183433,
final_local_scores=0.26180822, pem=-3.4974349,
similarity=30.922188]
20: 213137 James_Baker
    [pem_pos=0, final_scores=0.6091179,
final_local_scores=0.6009345, pem=0.0, similarity=22.022797]
21: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.3920228,
final_local_scores=0.40817314, pem=-0.1310191,
similarity=16.982916]
22: 11955 George_H._W._Bush
    [pem_pos=1, final_scores=0.15579961,
final_local_scores=0.107734025, pem=-0.80923575,
similarity=11.7054615]
```

Sample: b'901_Clinton'

[1 1 Clinton 1] wins [2 2 Democratic 2] re-nomination . [3 3 CHICAGO 3] 1996-08-28 President [4 4 Bill Clinton 4] was formally nominated on Wednesday as the [5 5 Democratic 5] party candidate for a second four-year term in the [6 6 White House 6] . [7 7 Clinton 7] won the nomination in a traditional state-by-state roll call of votes at the party convention and will accept in a speech on Thursday . He faces [8 8 Republican 8] challenger [9 9 Bob Dole 9] in the November 5 presidential election .

1: 3356 Bill_Clinton

```
2: 5043544 Democratic_Party_(United_States)
3: 6886 Chicago
4: 3356 Bill_Clinton
5: 5043544 Democratic_Party_(United_States)
6: 33057 White_House
7: 3356 Bill_Clinton
8: 32070 Republican_Party_(United_States)
9: 55846 Bob_Dole
1: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.78611296,
final_local_scores=0.75632423, pem=0.0, similarity=26.659937]
2: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.4661147,
final_local_scores=0.439412, pem=-0.076093495,
similarity=17.616453]
3: 6886 Chicago
    [pem_pos=0, final_scores=0.6036873,
final_local_scores=0.65018445, pem=-0.20658566,
similarity=24.61596]
4: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.62967235,
final_local_scores=0.5801703, pem=0.0, similarity=21.40315]
5: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.5040371,
final_local_scores=0.48083645, pem=-0.076093495,
similarity=18.85264]
6: 33057 White_House
    [pem_pos=0, final_scores=0.48373872,
final_local_scores=0.46645856, pem=-0.0056634876,
similarity=18.040565]
7: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.8701053,
final_local_scores=0.8480732, pem=0.0, similarity=29.39791]
8: 32070 Republican_Party_(United_States)
    [pem_pos=0, final_scores=0.45151037,
final_local_scores=0.42406136, pem=-0.1310191,
similarity=17.457052]
9: 55846 Bob_Dole
    [pem_pos=0, final_scores=0.7760443,
final_local_scores=0.7878498, pem=0.0, similarity=27.600721]
```

Sample: b'902_Beachcomber'

Beachcomber finds piece that could be [1 1 TWA 1] part . [2 2 ATLANTIC HIGHLANDS 2] , [3 3 N.J. 3] 1996-08-28 A foot-long piece of debris bearing markings from a commercial aircraft was found on the [4 4 New Jersey 4] shore and forwarded to [5 5 TWA 5] crash investigators in [6 6 Long Island 6] , officials said on Wednesday . A person walking on the shore at [7 7 Island Beach State Park 7] found the debris and alerted police who forwarded it to [8 8 Long Island 8] , [9 9 New York 9] , where the [10 10 National Transportation Safety Board 10] and the [11 11 FBI 11] are conducting an

investigation . The [12 12 TWA 12] jet exploded in a deadly fireball last month , killing 230 people , crashing in the Atlantic Ocean at least 80 miles from where the debris was found Wednesday . Several other items have been reported found along the [13 13 New Jersey 13] shore , most of it such personal items as wallets , shoes and jewelry . Investigators said they still do not have enough evidence to determine whether a bomb , a missile or mechanical failure caused the crash .

```
1: 152449 Trans_World_Airlines
2: 136974 Atlantic_Highlands,_New_Jersey
3: 21648 New_Jersey
4: 21648 New_Jersey
5: 152449 Trans_World_Airlines
6: 18315 Long_Island
7: 4472857 Island_Beach_State_Park
8: 18315 Long_Island
9: 8210131 New_York
10: 21840 National_Transportation_Safety_Board
11: 11127 Federal_Bureau_of_Investigation
12: 152449 Trans_World_Airlines
13: 21648 New_Jersey
1: 152449 Trans_World_Airlines
    [pem_pos=0, final_scores=0.5344997,
final_local_scores=0.5521872, pem=0.0, similarity=20.56808]
2: 136974 Atlantic_Highlands,_New_Jersey
    [pem_pos=0, final_scores=0.381831,
final_local_scores=0.4164247, pem=0.0, similarity=16.516657]
3: 21648 New_Jersey
    [pem_pos=0, final_scores=0.6075301,
final_local_scores=0.57142746, pem=0.0, similarity=21.142248]
4: 21648 New_Jersey
    [pem_pos=0, final_scores=0.43694884,
final_local_scores=0.38509333, pem=-0.13504621,
similarity=16.316069]
5: 152449 Trans_World_Airlines
    [pem_pos=0, final_scores=0.5447785,
final_local_scores=0.5634152, pem=0.0, similarity=20.903147]
6: 18315 Long_Island
    [pem_pos=0, final_scores=0.343992,
final_local_scores=0.29123998, pem=-0.1110181,
similarity=13.384628]
7: 4472857 Island_Beach_State_Park
    [pem_pos=0, final_scores=0.38154858,
final_local_scores=0.5113426, pem=0.0, similarity=19.349195]
8: 18315 Long_Island
    [pem_pos=0, final_scores=0.5194041,
final_local_scores=0.48285103, pem=-0.1110181,
similarity=19.102684]
9: 8210131 New_York
    [pem_pos=0, final_scores=0.35114664,
final_local_scores=0.30635384, pem=-0.2889045,
```

```
similarity=14.80303]
10: 21840 National_Transportation_Safety_Board
    [pem_pos=0, final_scores=0.39193374,
final_local_scores=0.45089066, pem=0.0, similarity=17.545189]
11: 11127 Federal_Bureau_of_Investigation
    [pem_pos=0, final_scores=0.4760995,
final_local_scores=0.53088516, pem=0.0, similarity=19.932384]
12: 152449 Trans_World_Airlines
    [pem_pos=0, final_scores=0.49323255,
final_local_scores=0.50710917, pem=0.0, similarity=19.222862]
13: 21648 New_Jersey
    [pem_pos=0, final_scores=0.43677825,
final_local_scores=0.384907, pem=-0.13504621,
similarity=16.310509]
```

Sample: b'903_FEATURE'

FEATURE - Rich new detail on [1 1 U.S. Civil War 1] unearthed . Leila Corcoran [2 2 WASHINGTON 2] 1996-08-29 A freed black man writes to his still-enslaved wife , a mother pleads with [3 3 Abraham Lincoln 3] on behalf of her son and a maimed soldier poses for an official photograph in newly reopened records that bring the [4 4 U.S. Civil War 4] back to life . Working in the basement of the National Archives , members of the Civil War Conservation Corps are organising the military records of volunteers who fought for the North so that they can be preserved on microfilm . The faded documents have been stowed away in the Archives since it opened in 1935 and have rarely seen the light of day . Each soldier 's file is a gold mine of information : enlistment papers , muster rolls , medical records , discharge certificates , letters and photographs . Since the project began almost two years ago , corps volunteers have focused on [5 5 African American 5] troops , preparing records in time for the unveiling in [6 6 Washington 6] this year of a special memorial to black soldiers who fought in the war . More than 185,000 black soldiers fought and 37,000 died . In one letter , a black soldier heading South wrote his wife , " though great is the present national difficulties yet I look forward to a brighter day when I shall have the oportunity of seeing you in the full enjoyment of freedom . " I would like to no (sic) if you are still in slavery if you are , it will not be long before we shall have crushed the system that now oppreses you for in the course of three months you shall have your liberty . Great is the outpouring of the coloured people that is now rallying with the heart of lions against the very curse that has separated you and me . " In another letter dated January 1865 , a well-to-do [7 7 Washington 7] matron wrote to [8 8 Lincoln 8] to plead for her son , who faced a dishonourable discharge from the [9 9 Army 9] . " James is a prisoner for a thoughtless act of folly , while those who have done nothing for the cause are free , " she wrote . [10 10 Lincoln 10] 's notation on the letter read : "

If his colonel will say in writing on this sheet he is willing to receive this man back to the regiment , I will pardon and send him . " The soldier was subsequently pardoned . While the letters speak to the anguish of separation , a photograph of a black amputee speaks to the terrible physical cost of the war . In a picture required for his military discharge , Pvt . Lewis Martin , with haunted eyes , posed bare-chested to reveal his missing arm and leg , blown off during a battle at [11 11 Petersburg 11] , [12 12 Virginia 12] , in July 1864 . The work on black troops has provided new insight into the rhythms of plantation talk and culture , said [13 13 John Simon 13] , professor of history at [14 14 Southern Illinois University at Carbondale 14] and editor of Gen . [15 15 Ulysses Grant 15] 's papers . " These (writings) are not only poetic but a linguistic treasure trove . This is the first generation of [16 16 African Americans 16] that really can express itself without fear and punishment , " [17 17 Simon 17] said . " I think there 's probably a whole lot of material that would expand our understanding of the sociology of the war , " said Edward Smith , director of [18 18 American Studies 18] at [19 19 American University 19] in [20 20 Washington 20] . The war between the North and [21 21 South 21] raged for nearly four years , claimed the lives of half a million [22 22 Americans 22] and forever seared the nation 's consciousness . 1 Southern records have already been preserved . They were put on microfilm about 30 years ago through a grant from the [23 23 United Daughters of the Confederacy 23] . The Civil War Conservation Corps is mostly retirees joined by students during the school year . Director Budge Weidman , who has shepherded the project from the beginning , predicts it will take up to a decade to complete . The work was inspired by a [24 24 National Park Service 24] plan to put computer databases at [25 25 Civil War 25] battlefields across the country so that [26 26 Americans 26] might research their ancestors . The records also provide insight into medical thinking of the day . One soldier was discharged because of " mental incapacity and inebetude (sic) of the brain , alleged by the Patient to be connected with a fall on the head but believed to arise from long continued and excessive masturbation . "

- 1: 863 American_Civil_War
- 2: 108956 Washington,_D.C.
- 3: 307 Abraham_Lincoln
- 4: 863 American_Civil_War
- 5: 2154 African_American
- 6: 108956 Washington,_D.C.
- 7: 108956 Washington,_D.C.
- 8: 307 Abraham_Lincoln
- 9: 32087 United_States_Army
- 10: 307 Abraham_Lincoln
- 11: 91268 Petersburg,_Virginia

```
12: 32432 Virginia
13: 27426216 John_Y._Simon
14: 476592 Southern_Illinois_University_Carbondale
15: 31752 Ulysses_S._Grant
16: 2154 African_American
17: 27426216 John_Y._Simon
18: 508127 American_studies
19: 80734 American_University
20: 108956 Washington,_D.C.
21: 7023 Confederate_States_of_America
22: 3434750 United_States
23: 178621 United_Daughters_of_the_Confederacy
24: 161535 National_Park_Service
25: 863 American_Civil_War
26: 3434750 United_States
1: 7023 Confederate_States_of_America
    [winner: 179553 Southern_United_States, pem_pos=0,
final_scores=-0.33974868, final_local_scores=-0.432226,
pem=-1.3130842, similarity=-1.6680168]
    RECALL MISS: 7023 Confederate_States_of_America]
1: 863 American_Civil_War
    [pem_pos=0, final_scores=0.40190095,
final_local_scores=0.35237098, pem=0.0, similarity=14.6051655]
2: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.59009385,
final_local_scores=0.56091803, pem=-1.7496481,
similarity=30.343483]
3: 307 Abraham_Lincoln
    [pem_pos=0, final_scores=0.46248287,
final_local_scores=0.41759008, pem=0.0, similarity=16.551434]
4: 863 American_Civil_War
    [pem_pos=0, final_scores=0.41618627,
final_local_scores=0.36797553, pem=0.0, similarity=15.070836]
5: 2154 African_American
    [pem_pos=1, final_scores=0.36145568,
final_local_scores=0.32440656, pem=-0.7252736,
similarity=17.714804]
6: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.5176826,
final_local_scores=0.4818198, pem=-1.7496481,
similarity=27.983034]
7: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.5192824,
final_local_scores=0.4835673, pem=-1.7496481,
similarity=28.035183]
8: 307 Abraham_Lincoln
    [pem_pos=0, final_scores=0.7338639,
final_local_scores=0.7173863, pem=0.0, similarity=25.497952]
9: 32087 United_States_Army
    [pem_pos=0, final_scores=0.3409838,
final_local_scores=0.3120495, pem=-1.138116,
similarity=19.591145]
```

```
10: 307 Abraham_Lincoln
    [pem_pos=0, final_scores=0.74168235,
final_local_scores=0.72592676, pem=0.0, similarity=25.752815]
11: 91268 Petersburg,_Virginia
    [pem_pos=0, final_scores=0.35547715,
final_local_scores=0.33624655, pem=-1.1939318,
similarity=20.616768]
12: 32432 Virginia
    [pem_pos=0, final_scores=0.6499792,
final_local_scores=0.635849, pem=-0.2023721,
similarity=24.165245]
13: 27426216 John_Y._Simon
    [pem_pos=8, final_scores=0.4483887,
final_local_scores=0.5419019, pem=-2.3025851,
similarity=32.782963]
14: 476592 Southern_Illinois_University_Carbondale
    [pem_pos=0, final_scores=0.34679645,
final_local_scores=0.3896922, pem=0.0, similarity=15.718905]
15: 31752 Ulysses_S._Grant
    [pem_pos=0, final_scores=0.5155665,
final_local_scores=0.49244577, pem=0.0, similarity=18.785276]
16: 2154 African_American
    [pem_pos=0, final_scores=0.4292624,
final_local_scores=0.3984751, pem=0.0, similarity=15.981006]
17: 27426216 John_Y._Simon
    [pem_pos=9, final_scores=0.5986982,
final_local_scores=0.70950764, pem=-2.3025851,
similarity=37.784653]
18: 508127 American_studies
    [pem_pos=0, final_scores=0.35554808,
final_local_scores=0.3989603, pem=0.0, similarity=15.995483]
19: 80734 American_University
    [pem_pos=0, final_scores=0.41835457,
final_local_scores=0.4349867, pem=0.0, similarity=17.070583]
20: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.5262956,
final_local_scores=0.49122816, pem=-1.7496481,
similarity=28.2638]
21: 7023 Confederate_States_of_America
    [pem_pos=19, final_scores=0.23539548,
final_local_scores=0.18845707, pem=-4.5094585,
similarity=34.236786]
22: 3434750 United_States
    [pem_pos=0, final_scores=0.36047548,
final_local_scores=0.34041768, pem=-0.5581965,
similarity=17.284014]
23: 178621 United_Daughters_of_the_Confederacy
    [pem_pos=0, final_scores=0.44723296,
final_local_scores=0.47625816, pem=0.0, similarity=18.302206]
24: 161535 National_Park_Service
    [pem_pos=0, final_scores=0.40538996,
final_local_scores=0.42362338, pem=0.0, similarity=16.73148]
```

```
25: 863 American_Civil_War
    [pem_pos=0, final_scores=0.4324538,
final_local_scores=0.38574535, pem=-0.48123205,
similarity=18.218138]
26: 3434750 United_States
    [pem_pos=0, final_scores=0.39657998,
final_local_scores=0.3798564, pem=-0.5581965,
similarity=18.460945]

Sample: b'904_Clinton'
[1 1 Clinton 1] arrives in [2 2 Chicago 2] on day of re-
nomination . [3 3 CHICAGO 3] 1996-08-28 President [4 4 Bill Cl
inton 4] arrived in [5 5 Chicago 5] on Wednesday as the [6 6 D
emocratic 6] convention prepared to re-nominate him for a
second four-year term . [7 7 Clinton 7] flew in by helicopter
from [8 8 Michigan City 8] , [9 9 Indiana 9] , after ending a
four-day , 559-mile trip aboard a campaign train from [10 10 W
ashington 10] .
1: 3356 Bill_Clinton
2: 6886 Chicago
3: 6886 Chicago
4: 3356 Bill_Clinton
5: 6886 Chicago
6: 5043544 Democratic_Party_(United_States)
7: 3356 Bill_Clinton
8: 112436 Michigan_City,_Indiana
9: 21883857 Indiana
10: 108956 Washington,_D.C.
1: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.7881989,
final_local_scores=0.7871643, pem=0.0, similarity=27.580265]
2: 6886 Chicago
    [pem_pos=0, final_scores=0.69623286,
final_local_scores=0.6878039, pem=-0.20658566,
similarity=25.738598]
3: 6886 Chicago
    [pem_pos=0, final_scores=0.6663353,
final_local_scores=0.65514535, pem=-0.20658566,
similarity=24.764002]
4: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.61951685,
final_local_scores=0.59282917, pem=0.0, similarity=21.780918]
5: 6886 Chicago
    [pem_pos=0, final_scores=0.6512322,
final_local_scores=0.6386475, pem=-0.20658566,
similarity=24.271671]
6: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.45144033,
final_local_scores=0.44982237, pem=-0.076093495,
similarity=17.927118]
7: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.8270989,
```

```
final_local_scores=0.82965654, pem=0.0, similarity=28.848318]
8: 112436 Michigan_City,_Indiana
    [pem_pos=0, final_scores=0.37171853,
final_local_scores=0.41526532, pem=0.0, similarity=16.482058]
9: 21883857 Indiana
    [pem_pos=0, final_scores=0.48529738,
final_local_scores=0.47706592, pem=-0.27230176,
similarity=19.80713]
10: 108956 Washington,_D.C.
    [pem_pos=0, final_scores=0.46996647,
final_local_scores=0.4274124, pem=-1.7496481,
similarity=26.35941]
```

Sample: b'905_New'

New bomb attacks on [1 1 Corsica 1] despite crackdown vow .
Sylvie Florence [2 2 AJACCIO 2] , [3 3 Corsica 3] 1996-08-29
Separatist guerrillas planted two bombs overnight at
government offices on the [4 4 French 4] [5 5 Mediterranean 5]
island of [6 6 Corsica 6] despite fresh warnings of a
crackdown by [7 7 Paris 7] , police said on Thursday . In the
latest in a wave of attacks , a two kg (four lb) bomb
seriously damaged two floors of Agriculture Ministry offices
located just 50 metres (yards) from a police station in the
centre of the island capital [8 8 Ajaccio 8] . No one was
hurt . A second device , packed with five kg (10 lbs) of
explosive , was defused before it could go off , police
said . The new attacks followed by a day a warning of a new "get-tough"
policy by [9 9 Paris 9] toward the separatists ,
who seek greater autonomy . Interior Minister [10 10 Jean-
Louis Debre 10] , under fire for staging secret talks with
one of the largest of several rival underground nationalist
groups , told the daily La Corse in a statement he had given
" firm orders " to police to round up those responsible for
the bombings and bring them to justice . Judges on the island
had accused [11 11 Paris 11] of taking a lax stance on
guerrilla violence while conducting secret but widely-
reported talks with separatists which have now failed . The
latest bombing , close on the heels of the new orders ,
brought charges that police were powerless . " No searches ,
no arrests , no police reinforcements visible on the island ,
despite the ministry 's promises , " the daily [12 12 France-
Soir 12] lamented . " On the island , as at the [13 13 Place B
eauvau 13] (the [14 14 Interior Ministry 14] 's [15 15 Paris
15] address) , people are well aware who is who and who is
doing what . It is time to end this nightly farce , " said
the pro-government daily [16 16 Le Figaro 16] in an
editorial . No one immediately claimed responsibility for
Thursday 's incidents , which brought to 23 the number of
guerrilla attacks on the resort island since mid-August ,
when separatist guerrillas ended a shaky seven-month truce . [
17 17 Corsica 17] has been racked by low-level separatist-
inspired violence , principally directed against government

targets , for two decades . The daily [18 18 Le Monde 18] reported on Wednesday some separatist movements were considering taking their attacks to the [19 19 French 19] mainland on the principle that " 300 grammes of explosives on the continent have more impact than 300 kilos in [20 20 Corsica 20] " . The newspaper said separatists may take advantage of social unrest widely expected on the mainland in coming weeks over government austerity plans to stoke a popular backlash against the government .

1: 5714828 Corsica
2: 2640 Ajaccio
3: 5714828 Corsica
4: 5843419 France
5: 19006 Mediterranean_Sea
6: 5714828 Corsica
7: 22989 Paris
8: 2640 Ajaccio
9: 22989 Paris
10: 1512904 Jean-Louis_Debré
11: 22989 Paris
12: 3017887 France_Soir
13: 17780202 Place_Beauvau
14: 871703 Minister_of_the_Interior_(France)
15: 22989 Paris
16: 80770 Le_Figaro
17: 5714828 Corsica
18: 127071 Le_Monde
19: 5843419 France
20: 5714828 Corsica
1: 5714828 Corsica
[pem_pos=0, final_scores=0.49144083,
final_local_scores=0.47337908, pem=-0.040366642,
similarity=18.43581]
2: 2640 Ajaccio
[pem_pos=0, final_scores=0.4691913,
final_local_scores=0.46444917, pem=0.0, similarity=17.949802]
3: 5714828 Corsica
[pem_pos=0, final_scores=0.42175967,
final_local_scores=0.397263, pem=-0.040366642,
similarity=16.164354]
4: 5843419 France
[pem_pos=0, final_scores=0.6913367,
final_local_scores=0.66833806, pem=-0.80223274,
similarity=28.396921]
5: 19006 Mediterranean_Sea
[pem_pos=0, final_scores=0.31156904,
final_local_scores=0.3370595, pem=-0.20345205,
similarity=15.254645]
6: 5714828 Corsica
[pem_pos=0, final_scores=0.48184198,
final_local_scores=0.46289378, pem=-0.040366642,
similarity=18.122908]

```
7: 22989 Paris
    [pem_pos=0, final_scores=0.61550564,
final_local_scores=0.59657127, pem=-0.115301244,
similarity=22.519615]
8: 2640 Ajaccio
    [pem_pos=0, final_scores=0.36403257,
final_local_scores=0.3495792, pem=0.0, similarity=14.521853]
9: 22989 Paris
    [pem_pos=0, final_scores=0.5347655,
final_local_scores=0.5083749, pem=-0.115301244,
similarity=19.887661]
10: 1512904 Jean-Louis_Debr 
    [pem_pos=0, final_scores=0.49756277,
final_local_scores=0.6174456, pem=0.0, similarity=22.51552]
11: 22989 Paris
    [pem_pos=0, final_scores=0.5808588,
final_local_scores=0.55872494, pem=-0.115301244,
similarity=21.390205]
12: 3017887 France_Soir
    [pem_pos=0, final_scores=0.5769103,
final_local_scores=0.6627771, pem=0.0, similarity=23.868305]
13: 17780202 Place_Beauvau
    [pem_pos=0, final_scores=0.33773327,
final_local_scores=0.4588207, pem=0.0, similarity=17.781837]
14: 871703 Minister_of_the_Interior_(France)
    [pem_pos=5, final_scores=0.20770131,
final_local_scores=0.22375166, pem=-2.3301098,
similarity=23.438408]
15: 22989 Paris
    [pem_pos=0, final_scores=0.6210381,
final_local_scores=0.60261464, pem=-0.115301244,
similarity=22.699963]
16: 80770 Le_Figaro
    [pem_pos=0, final_scores=0.49157196,
final_local_scores=0.52820224, pem=0.0, similarity=19.852322]
17: 5714828 Corsica
    [pem_pos=0, final_scores=0.44189996,
final_local_scores=0.41926324, pem=-0.040366642,
similarity=16.820885]
18: 127071 Le_Monde
    [pem_pos=0, final_scores=0.49080682,
final_local_scores=0.5433823, pem=0.0, similarity=20.305325]
19: 5843419 France
    [pem_pos=0, final_scores=0.7417136,
final_local_scores=0.72336715, pem=-0.80223274,
similarity=30.039097]
20: 5714828 Corsica
    [pem_pos=0, final_scores=0.46152383,
final_local_scores=0.44069928, pem=-0.040366642,
similarity=17.460579]
```

Sample: b'906_Sweden'

[1 1 Sweden 1] 's OM to open [2 2 London 2] forest products bourse . [3 3 STOCKHOLM 3] 1996-08-29 [4 4 Swedish 4] options and derivatives exchange OM Gruppen AB said on Thursday it would open an electronic bourse for forest industry products in [5 5 London 5] in the first half of 1997 . " Together with subsidiaries OMLX , the London Securities & Derivatives Exchange and OM Stockholm , OM Gruppen will open an international electronic bourse for forest products in the first half of 1997 , " OM Gruppen said in a statement . The first commodity to be traded on the PULPEX bourse will be pulp , but OM said trade would be extended to include products such as timber , recycled paper and other paper qualities . " Through the establishment of PULPEX , [6 6 Lond on 6] will have a commodities bourse for forest products which complements existing bourses for oil , metals and ' softs ' (coffee , sugar and cocoa) , " OM said . PULPEX was the result of a three-year project run in cooperation between OM and representatives of the forest industry , the company said . Huge swings in the price of pulp over the past few years have made pulp producers ' profitability unpredictable , and made investing in new production capacity a risky business . " Without the ability to hedge prices , changes in the world market price for pulp has had an immediate impact on players ' profitability , " OM said . It said global production of pulp amounted to around 200 million tonnes per year , of which around 20 percent or 40 million tonnes was sold on the spot market . At current prices , the value of this production was around \$ 25 billion . PULPEX will be both a marketplace and a clearing house , OM said , adding that the British Securities and Investments Board had been informed of OM 's plans . PULPEX 's clearing operation will be covered by the parent company guarantee issued by OM Gruppen to its wholly-owned bourses and clearing organisations . -- [7 7 Stockholm 7] newsroom , +46-8-700 1006

1: 5058739 Sweden
2: 17867 London
3: 26741 Stockholm
4: 5058739 Sweden
5: 17867 London
6: 17867 London
7: 26741 Stockholm

1: 5058739 Sweden
[pem_pos=0, final_scores=0.68965846,
final_local_scores=0.6749665, pem=-0.2135508,
similarity=25.393381]
2: 17867 London
[pem_pos=0, final_scores=0.6112824,
final_local_scores=0.5792539, pem=-0.09086737,
similarity=21.869955]
3: 26741 Stockholm
[pem_pos=0, final_scores=0.71839213,
final_local_scores=0.683089, pem=-0.1047564,

```
similarity=25.044134]
4: 5058739 Sweden
    [pem_pos=0, final_scores=0.6856124,
final_local_scores=0.67054677, pem=-0.63135606,
similarity=27.533579]
5: 17867 London
    [pem_pos=0, final_scores=0.5942877,
final_local_scores=0.56068975, pem=-0.09086737,
similarity=21.315964]
6: 17867 London
    [pem_pos=0, final_scores=0.47145325,
final_local_scores=0.42651188, pem=-0.09086737,
similarity=17.311829]
7: 26741 Stockholm
    [pem_pos=0, final_scores=0.66278696,
final_local_scores=0.6223488, pem=-0.1047564,
similarity=23.231524]
```

Sample: b'907_Amsterdam-Rotterdam-Antwerp'

Amsterdam-Rotterdam-Antwerp oil stock levels fall . [1 1 AMS TERDAM 1] 1996-08-29 Oil product inventories held in independent tankage in the Amsterdam-Rotterdam-Antwerp area were at the following levels , with week-ago and year-ago levels , industry sources said . All figures in thousands of tonnes : 29/8/96 22/8/96 1/9/95 Gasoline 400 400-425 425 Naphtha 50-75 75-100 50-75 Gas oil 1,600 1,650 1,850-1,900 Fuel oil 325 325-350 425 Jet kero 15 15-20 25 Motor gasoline stocks dipped slightly as barges left for [2 2 Germany 2] , but there were few inflows of cargoes . Naphtha inventories also dropped as [3 3 Germany 3] again took barges and no cargoes entered ARA . Gas oil stocks fell with some cargoes arriving from the former [4 4 Soviet Union 4] , but very fast throughput to markets in [5 5 Benelux 5] , [6 6 Germany 6] and [7 7 Switzerland 7] . Fuel oil inventories dipped slightly with some straight-run arrivals , but fair bunkering demand removing more material . Jet fuel stocks lowered as the aviation sector bought . -- Philip Blenkinsop , [8 8 Amsterdams 8] newsroom 31 20 504 5000

1: 844 Amsterdam

2: 11867 Germany

3: 11867 Germany

4: 26779 Soviet_Union

5: 4170 Benelux

6: 11867 Germany

7: 26748 Switzerland

8: 844 Amsterdam

1: 844 Amsterdam

```
    [pem_pos=0, final_scores=0.6503115,
final_local_scores=0.646301, pem=-0.0642803,
similarity=23.72619]
2: 11867 Germany
    [pem_pos=0, final_scores=0.7974706,
```

```
final_local_scores=0.7616976, pem=-0.15529114,
similarity=27.664785]
3: 11867 Germany
    [pem_pos=0, final_scores=0.8028999,
final_local_scores=0.76762825, pem=-0.15529114,
similarity=27.841768]
4: 26779 Soviet_Union
    [pem_pos=0, final_scores=0.3891148,
final_local_scores=0.409007, pem=-0.075323656,
similarity=16.70492]
5: 4170 Benelux
    [pem_pos=0, final_scores=0.41139072,
final_local_scores=0.43333912, pem=0.0, similarity=17.021416]
6: 11867 Germany
    [pem_pos=0, final_scores=0.86396486,
final_local_scores=0.8343325, pem=-0.15529114,
similarity=29.832357]
7: 26748 Switzerland
    [pem_pos=0, final_scores=0.62757146,
final_local_scores=0.6032577, pem=-0.18598008,
similarity=23.103514]
8: 844 Amsterdam
    [pem_pos=0, final_scores=0.65509826,
final_local_scores=0.6515297, pem=-0.0642803,
similarity=23.882227]
```

Sample: b'908_German'

[1 1 German 1] anti-nuclear activists in pantomime protest .
[2 2 BONN 2] 1996-08-29 About 200 [3 3 German 3] anti-nuclear
activists protested on Thursday against nuclear waste
transportation by re-enacting scenes from a demonstration
they staged in May that turned into a violent clash with
police . Activists dressed as police brandished batons and
firing a theatre-prop water cannon at " demonstrators " .
Police who had turned out in force all around the government
quarter fearing violence looked on in amusement . Last May
dozens of demonstrators and police were injured in violent
clashes around the [4 4 Gorleben 4] nuclear waste depot as
hundreds of protesters tried to block a delivery of waste by
train and truck .

```
1: 11867 Germany
2: 3295 Bonn
3: 11867 Germany
4: 4693197 Gorleben
1: 11867 Germany
    [pem_pos=1, final_scores=0.9248491,
final_local_scores=0.91784245, pem=-1.0465823,
similarity=37.17144]
2: 3295 Bonn
    [pem_pos=0, final_scores=0.48405367,
final_local_scores=0.4679739, pem=0.0, similarity=18.054987]
3: 11867 Germany
```

```
[pem_pos=1, final_scores=0.881517,
final_local_scores=0.8705086, pem=-1.0465823,
similarity=35.758904]
4: 4693197 Gorleben
[pem_pos=0, final_scores=0.31931233,
final_local_scores=0.40994847, pem=0.0, similarity=16.323393]

Sample: b'909_Italy'
[1 1 Italy 1] police arrest five over double Mafia killing .
[2 2 CATANIA 2] , [3 3 Sicily 3] 1996-08-29 [4 4 Italian 4]
Police said on Thursday they had arrested five people in
connection with the slaying of the daughter and 14-year-old
nephew of a Mafia boss earlier this week . Santa Puglisi ,
22 , and Salvatore Botta were gunned down in a cemetery in
this eastern [5 5 Sicilian 5] city on Tuesday in a crime
which shocked even hardened anti-Mafia investigators . The
arrests followed a tip-off from a married couple who
unexpectedly showed up at investigators ' offices on
Wednesday . " They wanted to get a weight off their
consciences , " a police spokesman said . He added the couple
had also shed light on the murder last year of the wife of
Mafia boss [6 6 Nitto Santapaola 6] . Puglisi was shot as she
knelt praying by the tomb of her young husband , who was
himself killed last year in a Mafia ambush . Botta , who had
accompanied Puglisi to the cemetery , tried to flee the lone
gunman but was caught and killed .
1: 14532 Italy
2: 44776 Catania
3: 27619 Sicily
4: 14532 Italy
5: 27619 Sicily
6: 8468733 Benedetto_Santapaola
1: 14532 Italy
[pem_pos=0, final_scores=0.7935372,
final_local_scores=0.77383524, pem=-0.20612453,
similarity=28.303438]
2: 44776 Catania
[pem_pos=0, final_scores=0.5582517,
final_local_scores=0.54465824, pem=-0.32144886,
similarity=22.09149]
3: 27619 Sicily
[pem_pos=0, final_scores=0.50304085,
final_local_scores=0.45983648, pem=-0.08753506,
similarity=18.28818]
4: 14532 Italy
[pem_pos=0, final_scores=0.79077727,
final_local_scores=0.77082044, pem=-0.75948656,
similarity=31.222736]
5: 27619 Sicily
[pem_pos=0, final_scores=0.5959191,
final_local_scores=0.56129175, pem=-0.66251653,
similarity=24.442646]
```

```
6: 8468733 Benedetto_Santapaola
    [pem_pos=0, final_scores=0.44287938,
final_local_scores=0.53699476, pem=0.0, similarity=20.114708]
```

Sample: b'910_Thieves'

Thieves make off with cash from prison canteen . [1 1 LIMERICK 1] , [2 2 Ireland 2] 1996-08-29 Thieves stole almost 2,000 [3 3 Irish 3] pounds (\$ 3,000) from the officers ' canteen of a [4 4 Limerick 4] jail on Thursday while warders slept in a room upstairs . Police said the thieves intercepted a woman arriving to work at the canteen , forced her to open the safe where takings were kept and made off with the cash . While the robbery was going on , several officers were asleep in a room over the canteen , which is in the grounds of the prison on [5 5 Ireland 5] 's southwest coast .

```
1: 56885 Limerick
```

```
2: 14560 Republic_of_Ireland
```

```
3: 14560 Republic_of_Ireland
```

```
4: 56885 Limerick
```

```
5: 14560 Republic_of_Ireland
```

```
1: 56885 Limerick
```

```
    [pem_pos=0, final_scores=0.4043975,
final_local_scores=0.35497382, pem=-0.62048864,
similarity=18.057154]
```

```
2: 14560 Republic_of_Ireland
```

```
    [pem_pos=1, final_scores=0.54905397,
final_local_scores=0.489927, pem=-1.5040284,
similarity=26.889252]
```

```
3: 14560 Republic_of_Ireland
```

```
    [pem_pos=3, final_scores=0.610125,
final_local_scores=0.5566378, pem=-2.2576945,
similarity=32.97859]
```

```
4: 56885 Limerick
```

```
    [pem_pos=0, final_scores=0.39388907,
final_local_scores=0.34349498, pem=-0.62048864,
similarity=17.714602]
```

```
5: 14560 Republic_of_Ireland
```

```
    [pem_pos=1, final_scores=0.6100896,
final_local_scores=0.5565992, pem=-1.5040284,
similarity=28.878885]
```

Sample: b'911_Italians'

[1 1 Italians 1] hold HIV-pensioner for harassing hookers . [2 2 GENOA 2] , [3 3 Italy 3] 1996-08-29 [4 4 Italian 4] police said on Thursday they had arrested a 61-year-old man after he fired blank shots at prostitutes he blamed for spreading [5 5 AIDS 5] . The pensioner , named only as Pietro T. , told investigators he was infected with HIV , the [6 6 AIDS 6] virus , and his wife had died of the disease . He did not say how they had contracted the illness . Police said the man had recently been spotted cruising red-light areas in this northern [7 7 Italian 7] city , hurling abuse at

prostitutes and firing blank shots at them . Police said they found in his apartment two fake guns and a pistol that would only fire blanks . " Although fake guns are legal , the use he made of them means he could be tried , " a police official said .

```
1: 14532 Italy
2: 66513 Genoa
3: 14532 Italy
4: 14532 Italy
5: 5069516 HIV/AIDS
6: 5069516 HIV/AIDS
7: 14532 Italy
1: 14532 Italy
    [pem_pos=1, final_scores=0.6351287,
final_local_scores=0.61059076, pem=-1.5574863,
similarity=30.78081]
2: 66513 Genoa
    [pem_pos=0, final_scores=0.5970234,
final_local_scores=0.62139475, pem=-0.40212637,
similarity=24.820198]
3: 14532 Italy
    [pem_pos=0, final_scores=0.7898471,
final_local_scores=0.7795971, pem=-0.20612453,
similarity=28.475384]
4: 14532 Italy
    [pem_pos=0, final_scores=0.78655916,
final_local_scores=0.77600557, pem=-0.75948656,
similarity=31.377472]
5: 5069516 HIV/AIDS
    [pem_pos=0, final_scores=0.5456116,
final_local_scores=0.5925707, pem=0.0, similarity=21.773205]
6: 5069516 HIV/AIDS
    [pem_pos=0, final_scores=0.47399604,
final_local_scores=0.5143416, pem=0.0, similarity=19.438692]
7: 14532 Italy
    [pem_pos=0, final_scores=0.7153267,
final_local_scores=0.698195, pem=-0.75948656,
similarity=29.055449]
```

Sample: b'912_Bodies'

Bodies found at site of [1 1 Russian 1] jet crash -
officials . [2 2 OSL0 2] 1996-08-29 Bodies have been sighted
but no survivors have yet been found at the site of Thursday
's crash of a [3 3 Russian 3] airliner on [4 4 Norway 4] 's
remote [5 5 Arctic 5] island of Spitzbergen , [6 6 Norwegian 6
] officials said . " We have found dead people , " said Rune
Hansen , the island 's deputy governor , told [7 7 Norwegian 7
] television . The [8 8 Norwegian 8] news agency NTB quoted
another official on the island as saying no survivors had
been found . The Vnukovo Airlines Tupolev 154 flight from [9 9
Moscow 9] , carrying 129 passengers and a crew of 12 ,
crashed in bad weather 10 km (six miles) east of [10 10 Lon

gyearbyen 10] , the island 's only airstrip , officials said . First rescuers arrived shortly after 1 p.m. (1100 [11 11 GMT 11]) and reported soon afterwards that most of the three-engine jet 's wreckage was scattered around the top of the small Opera mountain while the rest had slid down the mountainside . Air traffic officials said they had lost contact with the flight , scheduled to arrive at around 10.15 a.m. (0815 [12 12 GMT 12]) , shortly before it was due to land . Spitzbergen is a [13 13 Norwegian 13] coal-mining settlement . The only other community is in the [14 14 Russian 14] village of [15 15 Barentsburg 15] . [16 16 Russia 16] and [17 17 Norway 17] share the island 's resources under a treaty dating back to the 1920s .

1: 25391 Russia
2: 22309 Oslo
3: 25391 Russia
4: 21241 Norway
5: 36971 Arctic
6: 21241 Norway
7: 21241 Norway
8: 21241 Norway
9: 19004 Moscow
10: 337405 Longyearbyen
11: 12701 Greenwich_Mean_Time
12: 12701 Greenwich_Mean_Time
13: 21241 Norway
14: 25391 Russia
15: 466256 Barentsburg
16: 25391 Russia
17: 21241 Norway
1: 25391 Russia
[pem_pos=2, final_scores=0.71106714,
final_local_scores=0.71392286, pem=-1.0515022,
similarity=31.112825]
2: 22309 Oslo
[pem_pos=0, final_scores=0.5600002,
final_local_scores=0.56593823, pem=-0.035297178,
similarity=21.17039]
3: 25391 Russia
[pem_pos=1, final_scores=0.6081734,
final_local_scores=0.60152704, pem=-1.0515022,
similarity=27.75871]
4: 21241 Norway
[pem_pos=0, final_scores=0.751893,
final_local_scores=0.7309708, pem=-0.22668563,
similarity=27.136091]
5: 36971 Arctic
[pem_pos=0, final_scores=0.342866,
final_local_scores=0.37689555, pem=-0.102279626,
similarity=15.893242]
6: 21241 Norway
[pem_pos=0, final_scores=0.6073807,

```
final_local_scores=0.57311314, pem=-0.605967,
similarity=24.487896]
7: 21241 Norway
    [pem_pos=0, final_scores=0.59675235,
final_local_scores=0.5615033, pem=-0.605967,
similarity=24.141434]
8: 21241 Norway
    [pem_pos=0, final_scores=0.6484061,
final_local_scores=0.61792713, pem=-0.605967,
similarity=25.825233]
9: 19004 Moscow
    [pem_pos=0, final_scores=0.5686328,
final_local_scores=0.5832109, pem=-0.106582455,
similarity=22.073502]
10: 337405 Longyearbyen
    [pem_pos=0, final_scores=0.21578075,
final_local_scores=0.2278901, pem=0.0, similarity=10.8904085]
11: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.6611997,
final_local_scores=0.7520159, pem=0.0, similarity=26.531366]
12: 12701 Greenwich_Mean_Time
    [pem_pos=0, final_scores=0.6654858,
final_local_scores=0.75669783, pem=0.0, similarity=26.671085]
13: 21241 Norway
    [pem_pos=0, final_scores=0.6732295,
final_local_scores=0.645043, pem=-0.605967,
similarity=26.634426]
14: 25391 Russia
    [pem_pos=2, final_scores=0.78104454,
final_local_scores=0.7903625, pem=-1.0515022,
similarity=33.393936]
15: 466256 Barentsburg
    [pem_pos=0, final_scores=0.46390408,
final_local_scores=0.5261636, pem=0.0, similarity=19.791483]
16: 25391 Russia
    [pem_pos=0, final_scores=0.7358672,
final_local_scores=0.74101317, pem=-0.20299491,
similarity=27.306942]
17: 21241 Norway
    [pem_pos=0, final_scores=0.7057269,
final_local_scores=0.68054146, pem=-0.22668563,
similarity=25.63118]
```

Sample: b'913_Galeforce'
Galeforce winds and heavy rains batter [1 1 Belgium 1] . [2 2
BRUSSELS 2] 1996-08-29 Torrential rains and galeforce winds
battered [3 3 Belgium 3] on Thursday causing widespread
damage as some areas had more rainfall in 24 hours than they
normally get in a month , the meteorological office said .
Cellars were flooded , trees uprooted and roofs damaged , but
there were no reports of any injuries , an interior affairs
ministry spokesman said . Some trains were delayed as fallen

trees blocked lines . [4 4 Brussels 4] received 5.6 cm (2.24 inches) of water in the past 24 hours -- compared to an average 7.4 cm (2.96 inches) per month -- but in several communes in the south of the country up to 8 cm (3.2 inches) fell , the [5 5 Royal Meteorological Institute 5] (RMT) said . The RMT spokesman said that near the eastern city of [6 6 Turnhout 6] , a group of boy scouts camping in a low-lying meadow had to be evacuated as water flooded their tents . The rain also severely hindered [7 7 Belgian 7] investigators ' excavations in the southern village of [8 8 Ju met 8] , where they are looking for bodies in one of the houses of the main character in a paedophile sex-and-murder scandal . But the coastal towns got off lightly as the flooding that had been expected due to a combination of spring tides and high winds failed to materialise .

1: 3343 Belgium
2: 3708 Brussels
3: 3343 Belgium
4: 3708 Brussels
5: 16035441 Royal_Meteorological_Institute
6: 156847 Turnhout
7: 3343 Belgium
8: 17703237 Jumet
1: 3343 Belgium
[pem_pos=0, final_scores=0.73101795,
final_local_scores=0.69831324, pem=-0.24180195,
similarity=26.24373]
2: 3708 Brussels
[pem_pos=0, final_scores=0.60906535,
final_local_scores=0.5765878, pem=-0.025440974,
similarity=21.434595]
3: 3343 Belgium
[pem_pos=0, final_scores=0.56873125,
final_local_scores=0.5210397, pem=-0.24180195,
similarity=20.953535]
4: 3708 Brussels
[pem_pos=0, final_scores=0.5196514,
final_local_scores=0.47891665, pem=-0.025440974,
similarity=18.519894]
5: 16035441 Royal_Meteorological_Institute
[pem_pos=0, final_scores=0.37550604,
final_local_scores=0.5708644, pem=0.0, similarity=21.125444]
6: 156847 Turnhout
[pem_pos=0, final_scores=0.43443984,
final_local_scores=0.45091498, pem=0.0, similarity=17.545916]
7: 3343 Belgium
[pem_pos=0, final_scores=0.66401607,
final_local_scores=0.62512386, pem=-0.12403572,
similarity=23.41918]
8: 17703237 Jumet
[pem_pos=0, final_scores=0.3731144,
final_local_scores=0.5132754, pem=0.0, similarity=19.406874]

Sample: b'914_Repsol'

[1 1 Repsol 1] shares up 65 pesetas on H1 results . [2 2 MAD RID 2] 1996-08-29 Shares in [3 3 Spanish 3] oil and chemicals group [4 4 Repsol 4] were up 65 pesetas to 4,150 after the company announced net first half profits fell 1.1 percent to 61.45 billion pesetas on the previous year . This was close to the market 's forecast of net profits of 61.94 billion . Shares in [5 5 Repsol 5] shot up 100 pesetas to 4,175 shortly after the figures , having traded down 10 pesetas before the figures were released . " It 's madness , " said an analyst at a [6 6 Madrid 6] brokerage , adding that the market had overreacted to the news . " People were waiting for the results to come out before buying . The rise has been very quick and very crazy . " He predicted a correction , perhaps as early as this session . -- Madrid Newsroom +34 1 585 2161

1: 1076420 Repsol_S.A.

2: 41188263 Madrid

3: 26667 Spain

4: 1076420 Repsol_S.A.

5: 1076420 Repsol_S.A.

6: 41188263 Madrid

1: 1076420 Repsol_S.A.

[pem_pos=0, final_scores=0.5456759, final_local_scores=0.53861576, pem=0.0, similarity=20.163082]

2: 41188263 Madrid

[pem_pos=0, final_scores=0.5133781, final_local_scores=0.5235059, pem=-0.13236846, similarity=20.432014]

3: 26667 Spain

[pem_pos=1, final_scores=0.5424621, final_local_scores=0.57251817, pem=-1.2425895, similarity=27.932192]

4: 1076420 Repsol_S.A.

[pem_pos=0, final_scores=0.514899, final_local_scores=0.5049967, pem=0.0, similarity=19.159822]

5: 1076420 Repsol_S.A.

[pem_pos=0, final_scores=0.51774377, final_local_scores=0.5081042, pem=0.0, similarity=19.252556]

6: 41188263 Madrid

[pem_pos=0, final_scores=0.5358011, final_local_scores=0.5479997, pem=-0.13236846, similarity=21.162956]

Sample: b'915_Algeria'

[1 1 Algeria 1] forces kill guerrillas - papers . [2 2 PARIS 2] 1996-08-29 [3 3 Algerian 3] security forces killed four [4 4 Moslem 4] guerrillas on Tuesday in a village south of the capital [5 5 Algiers 5] , an [6 6 Algerian 6] newspaper said on Thursday . The armed militants were shot dead in Nahar village 90 km (56 miles) south of [7 7 Algiers 7] , Liberte newspaper said . Security forces also killed an unspecified

number of members of a rebel gang on Wednesday in the Leveilly suburb of [8 8 Algiers 8] , Le Matin newspaper reported . An estimated 50,000 people , mostly [9 9 Moslem 9] militants and security forces members , have been killed in violence pitting [10 10 Moslem 10] guerrillas against government forces since early 1992 , when authorities cancelled a general election in which [11 11 Islamists 11] had taken a commanding lead .

1: 358 Algeria
2: 22989 Paris
3: 358 Algeria
4: 6037917 Islam
5: 1644 Algiers
6: 358 Algeria
7: 1644 Algiers
8: 1644 Algiers
9: 6037917 Islam
10: 6037917 Islam
11: 15012 Islamism

1: 358 Algeria
 [pem_pos=0, final_scores=0.66765875,
final_local_scores=0.6578294, pem=-0.33694586,
similarity=25.553019]
2: 22989 Paris
 [pem_pos=0, final_scores=0.48758405,
final_local_scores=0.5147212, pem=-0.115301244,
similarity=20.077047]
3: 358 Algeria
 [pem_pos=0, final_scores=0.70567983,
final_local_scores=0.6993616, pem=-0.20917186,
similarity=26.097567]
4: 6037917 Islam
 [pem_pos=1, final_scores=0.46349645,
final_local_scores=0.4715181, pem=-1.2150909,
similarity=24.768606]
5: 1644 Algiers
 [pem_pos=0, final_scores=0.40986162,
final_local_scores=0.36969018, pem=-0.10460702,
similarity=15.690874]
6: 358 Algeria
 [pem_pos=0, final_scores=0.62731576,
final_local_scores=0.61376077, pem=-0.20917186,
similarity=23.543066]
7: 1644 Algiers
 [pem_pos=0, final_scores=0.40909606,
final_local_scores=0.368854, pem=-0.10460702,
similarity=15.665921]
8: 1644 Algiers
 [pem_pos=0, final_scores=0.386254,
final_local_scores=0.34390247, pem=-0.10460702,
similarity=14.921317]
9: 6037917 Islam

```
[pem_pos=1, final_scores=0.45870048,
final_local_scores=0.46627927, pem=-1.2150909,
similarity=24.612268]
10: 6037917 Islam
    [pem_pos=1, final_scores=0.45634216,
final_local_scores=0.4637031, pem=-1.2150909,
similarity=24.53539]
11: 15012 Islamism
    [pem_pos=0, final_scores=0.21084414,
final_local_scores=0.21916226, pem=0.0, similarity=10.6299515]
```

Sample: b'916_Finns'

```
[1 1 Finns 1] hold two men on child sex-abuse charges . [2 2
HELSINKI 2] 1996-08-29 [3 3 Finnish 3] police said on
Thursday they had arrested two men suspected of sexually
abusing a captive 13-year-old girl , but did not believe the
case was linked to others in [4 4 Europe 4] . The men , both [5
5 Finns 5] aged about 40 , were arrested last Saturday in
the western town of [6 6 Tampere 6] in a raid on a luxury
boat owned by one of them . The girl was being held on the
boat and had sought help from a passer-by , police chief
inspector Ilkka Laasonen said . " This is an individual case
and I do n't have any evidence linking the suspects to any
other cases , " Laasonen said . The girl was taken to
hospital . The men could face charges carrying up to six to
10 years in prison , he said . Witnesses had reported seeing
many young women and some girls who looked clearly underage
at parties around the boat in recent weeks , the daily
newspaper [7 7 Iltalehti 7] said .
```

```
1: 10577 Finland
2: 13696 Helsinki
3: 10577 Finland
4: 9239 Europe
5: 10577 Finland
6: 68836 Tampere
7: 2423598 Iltalehti
1: 10577 Finland
    [pem_pos=1, final_scores=0.5190264,
final_local_scores=0.45163983, pem=-1.7001518,
similarity=26.813236]
2: 13696 Helsinki
    [pem_pos=0, final_scores=0.5781844,
final_local_scores=0.5508813, pem=-0.015504186,
similarity=20.613426]
3: 10577 Finland
    [pem_pos=0, final_scores=0.7182118,
final_local_scores=0.6692196, pem=-0.8112885,
similarity=28.472473]
4: 9239 Europe
    [pem_pos=0, final_scores=0.6060243,
final_local_scores=0.65484923, pem=-0.086338006,
similarity=24.10124]
```

```
5: 10577 Finland
    [pem_pos=1, final_scores=0.5082148,
final_local_scores=0.43982977, pem=-1.7001518,
similarity=26.4608]
6: 68836 Tampere
    [pem_pos=0, final_scores=0.47231346,
final_local_scores=0.4420808, pem=0.0, similarity=17.282286]
7: 2423598 Iltalehti
    [pem_pos=0, final_scores=0.4102319,
final_local_scores=0.5150487, pem=0.0, similarity=19.459793]
```

Sample: b'917_Audi'

[1 1 Audi 1] CEO says expects no 96 currency impact . [2 2 IN GOLSTADT 2] 1996-08-29 [3 3 Audi AG 3] management board chairman Herbert Demel said on Thursday that the [4 4 German 4] luxury carmaker , a unit of [5 5 Volkswagen AG 5] , did not expect any burden on its 1996 results from currency market volatility . " We do not expect any burden on our 1996 results from currency markets , " Demel told [6 6 Reuters 6] in an interview . [7 7 Audi 7] would have been able to report a profit 300 million marks higher in 1995 if exchange rates had stayed the same as in 1994 , Demel had told a shareholder meeting last April . Demel said the carmaker had hedged about half of its currency risk for 1996 . -- John Gilardi , Frankfurt Newsroom , +49 69 756525

```
1: 848 Audi
2: 208251 Ingolstadt
3: 848 Audi
4: 11867 Germany
5: 32652 Volkswagen_Group
6: 18998750 Reuters
7: 848 Audi
1: 848 Audi
    [pem_pos=0, final_scores=0.37015057,
final_local_scores=0.31430832, pem=0.0, similarity=13.4693]
2: 208251 Ingolstadt
    [pem_pos=0, final_scores=0.29136074,
final_local_scores=0.32665256, pem=0.0, similarity=13.837677]
3: 848 Audi
    [pem_pos=0, final_scores=0.3819824,
final_local_scores=0.32723275, pem=0.0, similarity=13.854991]
4: 11867 Germany
    [pem_pos=1, final_scores=0.70227236,
final_local_scores=0.7634881, pem=-1.0465823,
similarity=32.5652]
5: 32652 Volkswagen_Group
    [pem_pos=0, final_scores=0.18736671,
final_local_scores=0.1635631, pem=0.0, similarity=8.970762]
6: 18998750 Reuters
    [pem_pos=0, final_scores=0.29044503,
final_local_scores=0.37510747, pem=0.0, similarity=15.283668]
7: 848 Audi
```

```
[pem_pos=0, final_scores=0.37820023,
final_local_scores=0.32310134, pem=0.0, similarity=13.731702]
```

Sample: b'918 KEKKILA'

KEKKILA SEES FULL-YR 1996 PROFIT VS LOSS . [1 1 HELSINKI 1] 1996-08-29 Fertilisers and saplings maker Kekkila Oy said on Thursday in a statement it expected a falling result trend in the latter half of the year , but a full-year profit was nevertheless likely . " Due to natural seasonal fluctuations in operations , the end-year result trend will be falling , but based on the early year result trend a profitable result is likely to be achieved , " Kekkila said in its January-June interim report . In 1995 , Kekkila reported a 5.6 million markka loss before extraordinary items and tax . In the first half , Kekkila posted a 6.1 million markka profit , up from 0.7 million . -- Helsinki Newsroom +358 - 0 - 680 50 245

1: 13696 Helsinki

1: 13696 Helsinki

```
[pem_pos=0, final_scores=0.3395763,
final_local_scores=0.6435227, pem=-0.015504186,
similarity=23.378029]
```

Sample: b'919 INTERVIEW-T&N'

INTERVIEW-T&N untroubled by margin pressure . [1 1 LONDON 1] 1996-08-29 The chairman of [2 2 British-based 2] components and engineering group T&N Plc said on Thursday the firm remained confident about the general prospects for its operating margins despite pressure from unsettled markets . In an interview following its first-half results , which included a less optimistic forecast for the second half of this year than it had made in the past , Sir Colin Hope said T&N had taken defensive action to protect it from patchy markets . Looking at market prospects , he said : " I think our best judgment at this stage is that it will probably bumble along in the rather mixed way it 's been in the first half . " " It 's very difficult to predict the market (trend) this year . It could be better , or it could be worse , " Hope added , echoing the demand uncertainty across automotive industries . " You can see anxieties in [3 3 Germany 3] and [4 4 France 4] , in particular , beginning to grow and develop . [5 5 America 5] , however , is looking a little better , " he said . Compared with the end of last year , when T&N predicted a sluggish first half and a rebound later in 1996 , Hope said : " I think the difference (now) is the first half has not actually been as bad as some felt it was going to be , but equally we 're certainly not predicting a recovery in the second half . " Against this background , Hope said the group was glad it had rationalised and destocked even though this had pressed margins , which had slipped to 9.5 percent from 11.3 percent a year ago . " I think the figure of 9.5 percent on the first half , when you consider the fact that we 've been destocking , the fairly

mixed customer demand , and the fact that we sold the (southern [6 6 African 6]) mines actually is not bad . " It does show how easily we should be able to bounce back over 10 percent again , " he added , saying : " We continue to feel very relaxed about our general view that we would average 10 percent profit margins over the cycle . " " We 've always taken the view that we are the sort of company that 's quite capable of working in difficult circumstances -- we 're rather used to it . And we feel very confident that we 're doing all the right things , " Hope said . " When the (profit) figures will bounce back up again is just a function of markets recovering just a fraction , " he added . Commenting on the continued struggle to get control of [7 7 German 7] piston maker Kolbenschmidt , which has been hampered by regulatory obstacles , Hope said T&N 's confidence " continues to improve " that it may eventually be able to proceed . -- Andrew Huddart , London Newsroom , +44 171 542 8716

```
1: 17867 London
2: 31717 United_Kingdom
3: 11867 Germany
4: 5843419 France
5: 3434750 United_States
6: 5334607 Africa
7: 11867 Germany
1: 17867 London
    [pem_pos=0, final_scores=0.69708914,
final_local_scores=0.7014629, pem=-0.09086737,
similarity=25.516916]
2: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.63231236,
final_local_scores=0.6045461, pem=0.0, similarity=22.130575]
3: 11867 Germany
    [pem_pos=0, final_scores=0.6848267,
final_local_scores=0.66151, pem=-0.15529114,
similarity=24.674988]
4: 5843419 France
    [pem_pos=0, final_scores=0.68600297,
final_local_scores=0.67370284, pem=-0.16240834,
similarity=25.07755]
5: 3434750 United_States
    [pem_pos=0, final_scores=0.5274522,
final_local_scores=0.49438912, pem=-0.78483206,
similarity=23.11131]
6: 5334607 Africa
    [pem_pos=0, final_scores=0.50284916,
final_local_scores=0.5235399, pem=-0.91165566,
similarity=24.670914]
7: 11867 Germany
    [pem_pos=1, final_scores=0.87941366,
final_local_scores=0.87406677, pem=-1.0465823,
similarity=35.865086]
```

Sample: b'920_Earthquake'

Earthquake jolts New Zealand [1 1 South Island 1] . [2 2 Wellington 2] 1996-08-29 An earthquake measuring 5.5 on the [3 3 Richter 3] scale shook New Zealand's upper [4 4 South Island 4] on Thursday but there were no reports of injuries , [5 5 Television New Zealand 5] said . It said the quake , centred near the small town of [6 6 Waiau 6] , was strongly felt in the cities of [7 7 Nelson 7] and [8 8 Christchurch 8] . Some minor damage had been reported in the spa town of Hanmer . [9 9 New Zealand 9] is prone to frequent earthquakes but they rarely cause major damage . The country has only 3.5 million people in an area about the size of [10 10 Britain 10] or [11 11 Japan 11] .

1: 62049 South_Island
2: 33804 Wellington
3: 19467352 Richter_magnitude_scale
4: 62049 South_Island
5: 369460 Television_New_Zealand
6: 2859334 Waiau,_Canterbury
7: 61978 Nelson,_New_Zealand
8: 628782 Christchurch
9: 4913064 New_Zealand
10: 31717 United_Kingdom
11: 15573 Japan
1: 62049 South_Island
 [pem_pos=0, final_scores=0.5501793,
final_local_scores=0.53759116, pem=0.0, similarity=20.132505]
2: 33804 Wellington
 [pem_pos=0, final_scores=0.47379822,
final_local_scores=0.4383782, pem=-0.9265605,
similarity=22.210573]
3: 19467352 Richter_magnitude_scale
 [pem_pos=0, final_scores=0.46715117,
final_local_scores=0.5710679, pem=-0.7852703,
similarity=25.401941]
4: 62049 South_Island
 [pem_pos=0, final_scores=0.49270207,
final_local_scores=0.47480595, pem=0.0, similarity=18.25887]
5: 369460 Television_New_Zealand
 [pem_pos=0, final_scores=0.5923703,
final_local_scores=0.6576949, pem=0.0, similarity=23.71664]
6: 2859334 Waiau,_Canterbury
 [pem_pos=2, final_scores=0.2818662,
final_local_scores=0.37869483, pem=-1.409942,
similarity=23.058205]
7: 61978 Nelson,_New_Zealand
 [pem_pos=0, final_scores=0.50137967,
final_local_scores=0.47919887, pem=-1.7804265,
similarity=28.072199]
8: 628782 Christchurch
 [pem_pos=0, final_scores=0.52050865,

```
final_local_scores=0.49311036, pem=-0.16020472,
similarity=19.676327]
9: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.5014163,
final_local_scores=0.4661852, pem=-0.23795462,
similarity=19.295643]
10: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5951447,
final_local_scores=0.6201846, pem=-0.61114657,
similarity=25.920769]
11: 15573 Japan
    [pem_pos=0, final_scores=0.67293745,
final_local_scores=0.7287685, pem=-0.13549899,
similarity=26.574484]
```

Sample: b'921_Hong'

[1 1 Hong Kong 1] 's [2 2 Tsang 2] sees growth , smooth transition . Mark Trevelyan [3 3 WELLINGTON 3] 1996-08-29 [4 4 Hong Kong 4] Financial Secretary [5 5 Donald Tsang 5] said on Thursday he expected the territory 's economy to keep growing at around five percent but with some fluctuations from year to year . [6 6 Tsang 6] , who made the remarks during a visit to [7 7 New Zealand 7] , also spoke strongly in favour of keeping the [8 8 Hong Kong 8] dollar pegged to its [9 9 U.S. 9] counterpart , and said negotiations with [10 10 China 10] on next year 's budget were going smoothly . [11 11 Hong Kong 11] 's economy grew by only 3.1 percent in the first quarter , down from 5.9 percent a year earlier , and some private sector economists have revised downwards their predictions for the 1996 year . Second quarter growth estimates will be released when the [12 12 Hong Kong 12] government issues its half-yearly economic report on Friday . " Our trend growth rate of five percent in real terms is pretty solid , " [13 13 Tsang 13] told a news conference after meeting [14 14 New Zealand 14] Finance Minister [15 15 Bill Birch 15] . " There will be fluctuations in individual years , but it wo n't be a big margin , " he said . He said inflation was under control and the [16 16 Hong Kong 16] dollar was " rock solid " . Its link to the [17 17 U.S. 17] dollar had proved an engine of growth for the past 12 years . " There 's absolutely no economic or financial or political reason for us to change . A lot of investment in [18 18 Hong Kong 18] , some of which (is) by [19 19 China 19] , is predicated on the link continuing . " [20 20 Britain 20] will hand over [21 21 Hong Kong 21] to [22 22 Chinese 22] sovereignty at midnight on June 30 , 1997 . [23 23 Tsang 23] said three sets of meetings with [24 24 Chinese 24] authorities on [25 25 Hong Kong 25] 's 1997-98 budget , which will span the transition period , had gone smoothly . " None of our basic precepts have been challenged . " [26 26 Hong Kong 26] will retain its own currency after the handover , run its own financial and monetary policy and have control over

its own foreign exchange reserves . It will have no duty to contribute any taxes to [27 27 Beijing 27] , [28 28 Tsang 28] said . He described the condition of the property market as " very good indeed " . " You know , we went through a little climb and a little trough over the last few years . Because of speculation in the market we introduced certain measures , but they are not draconian measures , and we brought it down to earth . " [29 29 Tsang 29] said the market would continue to appreciate because property and land were scarce . The government would put land on the market to stop rentals " going through the roof " , but this would mean reclamation , with possible environmental problems . " There will be on the whole a slightly upward climb , consistent with our economic growth rate , " he said in reference to the property market . [30 30 Tsang 30] , the third senior figure in the government after the governor and chief secretary , said his stated aim was to serve as financial secretary for two years under [31 31 British 31] rule and three years under [32 32 China 32] .

1: 13404 Hong_Kong
2: 336773 Donald_Tsang
3: 33804 Wellington
4: 13404 Hong_Kong
5: 336773 Donald_Tsang
6: 336773 Donald_Tsang
7: 4913064 New_Zealand
8: 13404 Hong_Kong
9: 3434750 United_States
10: 5405 China
11: 13404 Hong_Kong
12: 13404 Hong_Kong
13: 336773 Donald_Tsang
14: 4913064 New_Zealand
15: 595530 Bill_Birch
16: 13404 Hong_Kong
17: 3434750 United_States
18: 13404 Hong_Kong
19: 5405 China
20: 31717 United_Kingdom
21: 13404 Hong_Kong
22: 5405 China
23: 336773 Donald_Tsang
24: 5405 China
25: 13404 Hong_Kong
26: 13404 Hong_Kong
27: 18603746 Beijing
28: 336773 Donald_Tsang
29: 336773 Donald_Tsang
30: 336773 Donald_Tsang
31: 31717 United_Kingdom
32: 5405 China
1: 13404 Hong_Kong
[pem_pos=0, final_scores=0.60799384,

```
final_local_scores=0.5699219, pem=-0.22402541,
similarity=22.315605]
2: 336773 Donald_Tsang
    [pem_pos=0, final_scores=0.9489917,
final_local_scores=0.9709992, pem=0.0, similarity=33.066265]
3: 33804 Wellington
    [pem_pos=0, final_scores=0.4429255,
final_local_scores=0.47168887, pem=-0.9265605,
similarity=23.20463]
4: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.59289545,
final_local_scores=0.5534292, pem=-0.22402541,
similarity=21.82343]
5: 336773 Donald_Tsang
    [pem_pos=0, final_scores=0.76916295,
final_local_scores=0.7710151, pem=0.0, similarity=27.098341]
6: 336773 Donald_Tsang
    [pem_pos=0, final_scores=1.0026584,
final_local_scores=1.0296218, pem=0.0, similarity=34.81568]
7: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.58277935,
final_local_scores=0.61841524, pem=-0.23795462,
similarity=23.838491]
8: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.47954673,
final_local_scores=0.429613, pem=-0.22402541,
similarity=18.128508]
9: 3434750 United_States
    [pem_pos=0, final_scores=0.43597978,
final_local_scores=0.44260925, pem=-0.14075686,
similarity=18.063515]
10: 5405 China
    [pem_pos=0, final_scores=0.5402266,
final_local_scores=0.5365285, pem=-0.15669827,
similarity=20.952944]
11: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5523048,
final_local_scores=0.50909007, pem=-0.22402541,
similarity=20.500261]
12: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5108397,
final_local_scores=0.46379578, pem=-0.22402541,
similarity=19.14859]
13: 336773 Donald_Tsang
    [pem_pos=0, final_scores=1.0169872,
final_local_scores=1.045274, pem=0.0, similarity=35.282776]
14: 4913064 New_Zealand
    [pem_pos=0, final_scores=0.50633806,
final_local_scores=0.5349148, pem=-0.23795462,
similarity=21.346672]
15: 595530 Bill_Birch
    [pem_pos=0, final_scores=0.48400497,
```

```
final_local_scores=0.5974978, pem=0.0, similarity=21.92024]
16: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.4719956,
final_local_scores=0.4213645, pem=-0.22402541,
similarity=17.882355]
17: 3434750 United_States
    [pem_pos=0, final_scores=0.4837904,
final_local_scores=0.49483508, pem=-0.14075686,
similarity=19.622036]
18: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.48898882,
final_local_scores=0.43992704, pem=-0.22402541,
similarity=18.436298]
19: 5405 China
    [pem_pos=0, final_scores=0.49871546,
final_local_scores=0.49118394, pem=-0.15669827,
similarity=19.599771]
20: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.47548896,
final_local_scores=0.47663373, pem=-0.61114657,
similarity=21.636925]
21: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5175838,
final_local_scores=0.47116268, pem=-0.22402541,
similarity=19.368433]
22: 5405 China
    [pem_pos=1, final_scores=0.72776115,
final_local_scores=0.7413816, pem=-1.4196125,
similarity=33.934093]
23: 336773 Donald_Tsang
    [pem_pos=0, final_scores=1.0242667,
final_local_scores=1.0532258, pem=0.0, similarity=35.52007]
24: 5405 China
    [pem_pos=1, final_scores=0.69857913,
final_local_scores=0.70950466, pem=-1.4196125,
similarity=32.982822]
25: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5355584,
final_local_scores=0.49079716, pem=-0.22402541,
similarity=19.954365]
26: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.5606188,
final_local_scores=0.51817185, pem=-0.22402541,
similarity=20.77128]
27: 18603746 Beijing
    [pem_pos=0, final_scores=0.36183095,
final_local_scores=0.36761826, pem=-0.042161036,
similarity=15.289454]
28: 336773 Donald_Tsang
    [pem_pos=0, final_scores=0.92856115,
final_local_scores=0.9486819, pem=0.0, similarity=32.400272]
29: 336773 Donald_Tsang
```

```
[pem_pos=0, final_scores=1.0396062,
final_local_scores=1.0699818, pem=0.0, similarity=36.020107]
30: 336773 Donald_Tsang
    [pem_pos=0, final_scores=1.0369799,
final_local_scores=1.0671129, pem=0.0, similarity=35.93449]
31: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.55260783,
final_local_scores=0.5608744, pem=-0.48936236,
similarity=23.488552]
32: 5405 China
    [pem_pos=0, final_scores=0.51320666,
final_local_scores=0.5070134, pem=-0.15669827,
similarity=20.072153]
```

Sample: b'922_Taiwan'

```
[1 1 Taiwan 1] dollar ends higher , narrow trade seen . [2 2
TAIPEI 2] 1996-08-29 The [3 3 Taiwan 3] dollar closed
slightly firmer on Thursday amid tight [4 4 Taiwan 4] dollar
liquidity in the banking system , and dealers said the rate
was likely to move narrowly in the near term . The [5 5 Taiwan
5] dollar fell in early trade on month-end [6 6 U.S. 6]
dollar demand , but the downtrend was later reversed as [7 7 T
aiwan 7] dollar liquidity tightened . " Banks do not want to
hold big [8 8 U.S. 8] dollar positions at this moment ,
" said one dealer , adding that the rate was likely to hover
around current levels . The rate closed at T$ 27.482 against
Wednesday 's T$ 27.495 . Turnover was [9 9 US$ 9] 275
million . -- Joyce Liu ( 2-5080815 )
```

```
1: 25734 Taiwan
2: 57648 Taipei
3: 25734 Taiwan
4: 25734 Taiwan
5: 25734 Taiwan
6: 3434750 United_States
7: 25734 Taiwan
8: 3434750 United_States
9: 18717338 United_States_dollar
1: 25734 Taiwan
    [pem_pos=0, final_scores=0.53001493,
final_local_scores=0.46671414, pem=-0.077886656,
similarity=18.440954]
2: 57648 Taipei
    [pem_pos=0, final_scores=0.37606162,
final_local_scores=0.35618818, pem=0.0, similarity=14.719079]
3: 25734 Taiwan
    [pem_pos=0, final_scores=0.44285625,
final_local_scores=0.3715065, pem=-0.077886656,
similarity=15.599769]
4: 25734 Taiwan
    [pem_pos=0, final_scores=0.46106333,
final_local_scores=0.39139497, pem=-0.077886656,
similarity=16.19328]
```

```
5: 25734 Taiwan
    [pem_pos=0, final_scores=0.45838386,
final_local_scores=0.3884681, pem=-0.077886656,
similarity=16.105936]
6: 3434750 United_States
    [pem_pos=0, final_scores=0.5193186,
final_local_scores=0.5068267, pem=-0.14075686,
similarity=19.979889]
7: 25734 Taiwan
    [pem_pos=0, final_scores=0.4639361,
final_local_scores=0.39453304, pem=-0.077886656,
similarity=16.286926]
8: 3434750 United_States
    [pem_pos=0, final_scores=0.4784878,
final_local_scores=0.46222526, pem=-0.14075686,
similarity=18.648895]
9: 18717338 United_States_dollar
    [pem_pos=0, final_scores=0.44426614,
final_local_scores=0.4936543, pem=0.0, similarity=18.821342]
```

Sample: b'923_SIMEX'

SIMEX Nikkei ends down but off lows . [1 1 SINGAPORE 1]
1996-08-29 Simex Nikkei futures ended easier but off the day
's lows on Thursday . Dealers said selling in the session was
a follow-through from Wednesday 's gloomy Tankan corporate
report by the [2 2 Bank of Japan 2] . " The [3 3 Nikkei 3] is
testing support at the 20,500 level . Sentiment is a bit
gloomy because people are focusing on the weak recovery in
the economy at the moment , " said a dealer with a [4 4 Europ
ean 4] bank . September [5 5 Nikkei 5] settled at 20,605
after touching an intraday low of 20,530 against its previous
close of 20,725 . Volume was 19,560 contracts . Dealers said
technically , the index should see good support at 20,300 and
the upside should be capped at 21,000 . -- Doreen Siow

65-8703092

```
1: 27318 Singapore
2: 188155 Bank_of_Japan
3: 40293 Nikkei_225
4: 9239 Europe
5: 40293 Nikkei_225
1: 27318 Singapore
    [pem_pos=0, final_scores=0.43787503,
final_local_scores=0.5095503, pem=-0.1922004,
similarity=20.340925]
2: 188155 Bank_of_Japan
    [pem_pos=0, final_scores=0.6173831,
final_local_scores=0.711217, pem=0.0, similarity=25.313847]
3: 40293 Nikkei_225
    [pem_pos=1, final_scores=0.43715817,
final_local_scores=0.4528424, pem=-0.46501607,
similarity=20.132263]
4: 9239 Europe
```

```
[pem_pos=0, final_scores=0.30471426,
final_local_scores=0.40150958, pem=-0.25557294,
similarity=17.461405]
5: 40293 Nikkei_225
    [pem_pos=1, final_scores=0.40817815,
final_local_scores=0.42118615, pem=-0.46501607,
similarity=19.187578]

Sample: b'924_Siam'
Siam Commercial wins agency bond auctions . [1 1 BANGKOK 1]
1996-08-29 A consortium led by [2 2 Thailand 2] 's Siam
Commercial Bank Plc has secured at auction the right to sell
two state agency bond issues worth a combined 3.73 billion
baht , an official at the bank said on Thursday . The
Government Housing Bank will issue bonds worth three billion
baht and the metropolitan Waterworks Authority will issue
bonds worth 730 million , an investment banker at [3 3 Siam Co
mmercial Bank 3] told [4 4 Reuters 4] . The consortium , made
up of eight financial institutions , offered an annual
interest rate of 8.46 percent for both issues , he said .
Both state agency bonds will have seven-year maturity and
will be issued on September 5 , he said . -- [5 5 Bangkok 5]
newsroom ( 662 ) 652-0642
1: 56615 Bangkok
2: 30128 Thailand
3: 18143347 Siam_Commercial_Bank
4: 18998750 Reuters
5: 56615 Bangkok
1: 56615 Bangkok
    [pem_pos=0, final_scores=0.5529409,
final_local_scores=0.51866114, pem=-0.0005963434,
similarity=19.570839]
2: 30128 Thailand
    [pem_pos=0, final_scores=0.7121456,
final_local_scores=0.7286696, pem=-0.24029273,
similarity=27.141415]
3: 18143347 Siam_Commercial_Bank
    [pem_pos=0, final_scores=0.4406222,
final_local_scores=0.5626967, pem=0.0, similarity=20.881704]
4: 18998750 Reuters
    [pem_pos=0, final_scores=0.359429,
final_local_scores=0.44517362, pem=0.0, similarity=17.374582]
5: 56615 Bangkok
    [pem_pos=0, final_scores=0.62165,
final_local_scores=0.59371537, pem=-0.0005963434,
similarity=21.810606]
```

```
Sample: b"925_M'bishi"
M'bishi Gas sets terms on 7-year straight . [1 1 TOKYO 1]
1996-08-29 BORROWER - Mitsubishi Gas Chemical Co Ltd LEAD MGR
- Nomura Securities Co Ltd FISCAL AGENT - [2 2 Tokyo-
Mitsubishi Bank 2] TYPE straight bond ISSUE NO 13 AMT 10 bln
```

yen COUPON 2.95 % ISS PRICE MATURITY 5 . Sep.03 LAST MOODY 'S
PAY DATE 5 . Sep.96 FIRST INT PAY 5 . Mar.97 INT PAY 5 .
Mar / Sep LAST [3 3 S&P 3] SIGN DATE SUB DATE 5 . Jul-18.Jul
LAST JCR LAST JBRI A LAST NIS
1: 30057 Tokyo
2: 2823064 The_Bank_of_Tokyo-Mitsubishi_UFJ
3: 164386 Standard_&_Poor's
1: 30057 Tokyo
[pem_pos=0, final_scores=0.5075794,
final_local_scores=0.6306964, pem=-0.09283106,
similarity=23.415781]
2: 2823064 The_Bank_of_Tokyo-Mitsubishi_UFJ
[pem_pos=0, final_scores=0.39775276,
final_local_scores=0.45311666, pem=0.0, similarity=17.611618]
3: 164386 Standard_&_Poor's
[pem_pos=0, final_scores=0.5875789,
final_local_scores=0.70382905, pem=0.0, similarity=25.093376]

Sample: b'926_S.'
[1 1 S. Korea 1] [2 2 Daewoo 2] , [3 3 Dacom 3] units in [4 4
Polish 4] telecom JV . [5 5 SEOUL 5] 1996-08-29 [6 6 South Ko
rea 6] 's Daewoo Corp , unlisted Daewoo Information Systems
Co Ltd , Dacom Corp and Dacom International have set up a
joint venture to offer telecommunications services in [7 7 Po
land 7] . Daewoo Dacom Communications ([8 8 Poland 8]) [9 9
Ltd 9] , which was set up with an initial investment of \$ 1.0
million , is expected to have sales of \$ 60 million by the
year 2000 , a [10 10 Daewoo 10] statement said on Thursday ..
Daewoo Corp will take a 31 percent stake in the venture ,
Dacom International 25 percent , [11 11 Dacom 11] 24
percent , and Daewoo Information 20 percent , the statement
said . Daewoo Corp and Daewoo Information are units of [12 12
Daewoo Group 12] . -- Seoul Newsroom (822) 727 5644
1: 27019 South_Korea
2: 306128 Daewoo
3: 3109012 LG_Uplus
4: 22936 Poland
5: 19159283 Seoul
6: 27019 South_Korea
7: 22936 Poland
8: 22936 Poland
9: 538325 Limited_company
10: 306128 Daewoo
11: 3109012 LG_Uplus
12: 306128 Daewoo
1: 27019 South_Korea
[pem_pos=0, final_scores=0.5327442,
final_local_scores=0.51505065, pem=0.0, similarity=19.459852]
2: 306128 Daewoo
[pem_pos=0, final_scores=0.27767986,
final_local_scores=0.28322625, pem=-0.04748799,
similarity=12.799996]

```
3: 3109012 LG_Uplus
    [pem_pos=0, final_scores=0.2881286,
final_local_scores=0.34712642, pem=0.0, similarity=14.448658]
4: 22936 Poland
    [pem_pos=0, final_scores=0.58751005,
final_local_scores=0.62523514, pem=-0.81518245,
similarity=27.181065]
5: 19159283 Seoul
    [pem_pos=0, final_scores=0.4814564,
final_local_scores=0.4849655, pem=-0.0378014,
similarity=18.767622]
6: 27019 South_Korea
    [pem_pos=0, final_scores=0.58565813,
final_local_scores=0.572851, pem=-0.22449642,
similarity=22.405577]
7: 22936 Poland
    [pem_pos=0, final_scores=0.52315,
final_local_scores=0.55493164, pem=-0.15333635,
similarity=21.483847]
8: 22936 Poland
    [pem_pos=0, final_scores=0.46057492,
final_local_scores=0.48657787, pem=-0.15333635,
similarity=19.444033]
9: 538325 Limited_company
    [pem_pos=0, final_scores=0.46000093,
final_local_scores=0.53621984, pem=-0.403316,
similarity=22.284878]
10: 306128 Daewoo
    [pem_pos=0, final_scores=0.2628036,
final_local_scores=0.2669762, pem=-0.04748799,
similarity=12.3150625]
11: 3109012 LG_Uplus
    [pem_pos=0, final_scores=0.39481837,
final_local_scores=0.46366882, pem=0.0, similarity=17.926516]
12: 306128 Daewoo
    [pem_pos=0, final_scores=0.44199342,
final_local_scores=0.46271384, pem=0.0, similarity=17.898016]
```

Sample: b'927_Salomon'

Salomon & Taylor - 96/97 div forecast . [1 1 TOKYO 1]
1996-08-29 Year to March 31 , 1997 (in billions of yen
unless specified) LATEST ACTUAL (Parent) FORECAST YEAR-AGO
Ord div 10.00 yen 8.00 yen - Commem div - 2.00 yen NOTE -
Salomon & Taylor Made Co Ltd manufactures golf clubs and
sells ski equipment .

1: 30057 Tokyo

1: 30057 Tokyo

```
    [pem_pos=0, final_scores=0.43835413,
final_local_scores=0.75142246, pem=-0.09283106,
similarity=27.018488]
```

Sample: b'928_Indonesia'

[1 1 Indonesia 1] plays down [2 2 U.S. 2] consulate attack .
[3 3 JAKARTA 3] 1996-08-29 [4 4 Indonesia 4] sought on Thursday to play down a fire bomb attack on a [5 5 U.S. 5] consulate earlier this week , saying it was being treated as a criminal rather than a political act , the official [6 6 An tara 6] news agency said . The police are still investigating the incident . It is evident the happening is not politically motivated but just an ordinary criminal act , the news agency quoted [7 7 East Java 7] military commander Major-General [8 8 Utomo 8] as saying . The Tuesday morning attack on the consulate in Indonesias second largest city of [9 9 Surabaya 9] , caused slight damage to a guard house before being quickly extinguished , a spokesman at the [10 10 U.S. 10] embassy in [11 11 Jakarta 11] , 700 km (430 miles) west of [12 12 Surabaya 12] , said on Wednesday . No one was injures , he said . Nobody should try and exaggerate it by calling it a bomb because it was just a molotov cocktail , [13 13 Utomo 13] said . He said police were investigating the incident and patrols around diplomatic offices in [14 14 Surabaya 14] would be stepped up .

1: 14579 Indonesia
2: 3434750 United_States
3: 16275 Jakarta
4: 14579 Indonesia
5: 3434750 United_States
6: 2917433 Antara_(news_agency)
7: 320443 East_Java
8: 12360951 Imam_Utomo
9: 212635 Surabaya
10: 3434750 United_States
11: 16275 Jakarta
12: 212635 Surabaya
13: 12360951 Imam_Utomo
14: 212635 Surabaya
1: 14579 Indonesia
 [pem_pos=0, final_scores=0.6096496,
final_local_scores=0.5880562, pem=-0.23151328,
similarity=22.89749]
2: 3434750 United_States
 [pem_pos=0, final_scores=0.49258262,
final_local_scores=0.52190703, pem=-0.14075686,
similarity=20.429918]
3: 16275 Jakarta
 [pem_pos=0, final_scores=0.5405115,
final_local_scores=0.49274856, pem=0.0, similarity=18.794313]
4: 14579 Indonesia
 [pem_pos=0, final_scores=0.5914089,
final_local_scores=0.5681311, pem=-0.23151328,
similarity=22.302883]
5: 3434750 United_States
 [pem_pos=0, final_scores=0.4711706,
final_local_scores=0.4985177, pem=-0.14075686,

```
similarity=19.731932]
6: 2917433 Antara_(news_agency)
    [pem_pos=1, final_scores=0.4957776,
final_local_scores=0.5641474, pem=-1.1294895,
similarity=27.067337]
7: 320443 East_Java
    [pem_pos=0, final_scores=0.37795132,
final_local_scores=0.35001653, pem=0.0, similarity=14.5349045]
8: 12360951 Imam_Utomo
    [pem_pos=0, final_scores=0.49473488,
final_local_scores=0.6469341, pem=0.0, similarity=23.395517]
9: 212635 Surabaya
    [pem_pos=0, final_scores=0.46856773,
final_local_scores=0.43471587, pem=0.0, similarity=17.062502]
10: 3434750 United_States
    [pem_pos=0, final_scores=0.5018099,
final_local_scores=0.5319864, pem=-0.14075686,
similarity=20.730705]
11: 16275 Jakarta
    [pem_pos=0, final_scores=0.45218718,
final_local_scores=0.39626765, pem=0.0, similarity=15.91513]
12: 212635 Surabaya
    [pem_pos=0, final_scores=0.4444213,
final_local_scores=0.40833956, pem=0.0, similarity=16.27538]
13: 12360951 Imam_Utomo
    [pem_pos=0, final_scores=0.40555477,
final_local_scores=0.54951835, pem=0.0, similarity=20.488436]
14: 212635 Surabaya
    [pem_pos=0, final_scores=0.37500483,
final_local_scores=0.33251262, pem=0.0, similarity=14.012552]
```

Sample: b'929_Airport'

Airport Facilities - 6mth parent forecast . [1 1 TOKYO 1]
1996-08-29 Six months to September 30 , 1996 (in billions of
yen unless specified) LATEST PREVIOUS ACTUAL (Parent)
FORECAST FORECAST YEAR-AGO Sales 11.38 11.38 11.45 Current
1.09 1.09 918 million Net 934 million 490 million 538 million
NOTE - Airport Facilities Co Ltd manages and rents facilities
at [2 2 Haneda 2] ([3 3 Tokyo 3]) and [4 4 Itami 4] ([5 5 O
saka 5]) airports .

```
1: 30057 Tokyo
2: 24071884 Haneda_Airport
3: 30057 Tokyo
4: 24086495 Osaka_International_Airport
5: 52230 Osaka
1: 30057 Tokyo
    [pem_pos=0, final_scores=0.73041654,
final_local_scores=0.72355133, pem=-0.09283106,
similarity=26.186758]
2: 24071884 Haneda_Airport
    [pem_pos=0, final_scores=0.3814066,
final_local_scores=0.37705135, pem=0.0, similarity=15.341678]
```

```
3: 30057 Tokyo
    [pem_pos=0, final_scores=0.47235817,
final_local_scores=0.44166178, pem=-0.09283106,
similarity=17.77461]
4: 24086495 Osaka_International_Airport
    [pem_pos=1, final_scores=0.2910443,
final_local_scores=0.32458478, pem=-1.0835332,
similarity=19.668392]
5: 52230 Osaka
    [pem_pos=0, final_scores=0.3678761,
final_local_scores=0.3188924, pem=-0.08992924,
similarity=14.095147]
```

Sample: b'930_HK'

HK civil servants contest ban on [1 1 China 1] panel . [2 2 H
ONG KONG 2] 1996-08-29 Senior [3 3 Hong Kong 3] civil
servants were given the go-ahead on Thursday to challenge a
government ban on them standing for the Beijing-backed panel
to choose the territory 's first post-handover leader and
lawmakers . The 1 Supreme Court ruled that a judicial hearing
contesting the ban would be heard on September 11 , three
days before the nomination period for the Selection Committee
closes . The government maintains the ban , announced earlier
this month , is necessary to avoid a possible conflict of
interest because civil servants are involved in determining
government policy . Civil servants argue the ban stymies
their political rights . The 400-strong Selection Committee
will select [4 4 Hong Kong 4] 's future chief executive to
replace the [5 5 British 5] governor and a provisional
legislature to take over from the elected chamber which [6 6 B
eijing 6] plans to dissolve . [7 7 Hong Kong 7] , a [8 8 Brit
ish 8] colony for more than 150 years , will be handed back
to [9 9 China 9] at midnight on June 30 next year . [10 10 Ch
ina 10] intends to dismantle the territory 's first fully-
elected legislature because it opposes [11 11 Britain 11] 's
recent electoral reforms and install an interim appointed
chamber , a decision that has generated considerable
controversy . The judicial review sought by directorate-grade
bureaucrats will apply to only about 1,000 of the
approximately 33,000 civil servants affected . Police unions
are not contesting the ban , which affects all 27,000
officers , and nor are the very top tier of [12 12 Hong Kong 1
2] 's mandarin class , the policy secretaries . More than
16,000 application forms for places on the Selection
Committee have been handed out since the nomination period
opened . It closes on September 14 .

```
1: 5405 China
2: 13404 Hong_Kong
3: 13404 Hong_Kong
4: 13404 Hong_Kong
5: 31717 United_Kingdom
6: 18603746 Beijing
```

7: 13404 Hong_Kong
8: 31717 United_Kingdom
9: 5405 China
10: 5405 China
11: 31717 United_Kingdom
12: 13404 Hong_Kong
1: 7097642 Supreme_Court_of_the_Republic_of_China
[winner: 31737 Supreme_Court_of_the_United_States,
pem_pos=0, final_scores=-0.08985703,
final_local_scores=-0.036881313, pem=-0.9940822,
similarity=8.395075]
RECALL MISS: 7097642
Supreme_Court_of_the_Republic_of_China]
1: 5405 China
[pem_pos=0, final_scores=0.56460816,
final_local_scores=0.53433007, pem=-0.15669827,
similarity=20.887339]
2: 13404 Hong_Kong
[pem_pos=0, final_scores=0.6309086,
final_local_scores=0.5801443, pem=-0.22402541,
similarity=22.62066]
3: 13404 Hong_Kong
[pem_pos=0, final_scores=0.5454181,
final_local_scores=0.48675883, pem=-0.22402541,
similarity=19.833853]
4: 13404 Hong_Kong
[pem_pos=0, final_scores=0.5442636,
final_local_scores=0.48549777, pem=-0.22402541,
similarity=19.79622]
5: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5394236,
final_local_scores=0.51020485, pem=-0.48936236,
similarity=21.976473]
6: 18603746 Beijing
[pem_pos=0, final_scores=0.31325632,
final_local_scores=0.28915754, pem=-0.042161036,
similarity=12.948029]
7: 13404 Hong_Kong
[pem_pos=0, final_scores=0.514555,
final_local_scores=0.4530456, pem=-0.22402541,
similarity=18.827784]
8: 31717 United_Kingdom
[pem_pos=0, final_scores=0.5769401,
final_local_scores=0.551186, pem=-0.48936236,
similarity=23.199432]
9: 5405 China
[pem_pos=0, final_scores=0.51931566,
final_local_scores=0.48485488, pem=-0.15669827,
similarity=19.4109]
10: 5405 China
[pem_pos=0, final_scores=0.51495236,
final_local_scores=0.48008865, pem=-0.15669827,

```
similarity=19.268665]
11: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5589163,
final_local_scores=0.53149766, pem=-0.61114657,
similarity=23.274174]
12: 13404 Hong_Kong
    [pem_pos=0, final_scores=0.57492685,
final_local_scores=0.5189927, pem=-0.22402541,
similarity=20.795776]
```

Sample: b'931_Singapore'

Singapore Refining Company expected to shut CDU 3 . [1 1 SIN
GAPORE 1] 1996-08-29 Singapore Refining Company (SRC) is
expected to shutdown its 60,000 barrel-per-day (bpd) crude
distillation unit (CDU) in September , an industry source
said on Thursday . " They think something is stuck , " the
source said . " But nothing has been decided as they are
still waiting for an X-ray machine to determine the problem .
" The source said the problem was discovered during the past
month during which time CDU No.3 's production has varied
from maximum capacity of 60,000 bpd to as low as 40,000 bpd ,
depending on the crude being run . " We are having a lot of
problems , " said a source . An earlier problem with CDU No .
3 arose on July 24 when an industry source said the CDU would
be closed down briefly in August for repairs to a heat
exchanger . But by mid-August , a company spokesman said
repairs had been carried out without any shutdown . The
285,000 bpd SRC refinery is co-owned by the Singapore
Petroleum Company , British Petroleum and [2 2 Caltex 2] , the
joint-venture of [3 3 U.S. 3] majors [4 4 Chevron Corp 4] and
Texaco Inc . -- Singapore Newsroom (+65-8703086)

```
1: 27318 Singapore
2: 3467272 Caltex
3: 3434750 United_States
4: 284749 Chevron_Corporation
1: 27318 Singapore
    [pem_pos=0, final_scores=0.53039724,
final_local_scores=0.5509367, pem=-0.1922004,
similarity=21.57598]
2: 3467272 Caltex
    [pem_pos=0, final_scores=0.21200936,
final_local_scores=0.23126793, pem=0.0, similarity=10.991209]
3: 3434750 United_States
    [pem_pos=0, final_scores=0.53876656,
final_local_scores=0.57985467, pem=-0.14075686,
similarity=22.159191]
4: 284749 Chevron_Corporation
    [pem_pos=0, final_scores=0.47939473,
final_local_scores=0.5066717, pem=0.0, similarity=19.209808]
```

Sample: b'932_Loxley'

Loxley H1 net rises to 332.66 mln baht . [1 1 BANGKOK 1]

1996-08-29 Reviewed financial results for the first six months ended June 30 , 1996 . (in millions of baht unless stated) Six months 1996 1995 Shr (baht) 8.32 vs 6.66 Net 332.66 vs 266.37 NOTES : Second quarter figures not available . Full name of company is Loxley Publications Plc . -- [2 2 Bangkok 2] newsroom 662-252-9950

1: 56615 Bangkok

2: 56615 Bangkok

1: 56615 Bangkok

[pem_pos=0, final_scores=0.5774928,
final_local_scores=0.50273174, pem=-0.0005963434,
similarity=19.095474]

2: 56615 Bangkok

[pem_pos=0, final_scores=0.6703504,
final_local_scores=0.6041645, pem=-0.0005963434,
similarity=22.122429]

Sample: b'933_INDICATORS'

INDICATORS - [1 1 Spain 1] - updated August 29 . INDICATORS - monthly MTH / MTH PVS YR-AGO INDEX TOTAL CPI (%) Jul +0.1 - 0.1 +0.0 119.3** - Yr / yr Inflation (%) +3.7 +3.6 +4.7 119.3 - Core Inflation +0.1 +0.2 +0.2 - - Yr / yr rise +3.5 +3.6 +5.2 - - JOBLESS (INEM) Jul - 63,913 - 33,149 - 65,345 - 2.17M Rate (%) 13.67 14.15 15.19 - - BALANCE OF PAYMENTS Trade (bln pts) May - 196.8 - 180.6 - 279.9 - - Cur Acc (bln pts) May - 9.5 - 42.0 - 110.4 - - RESERVES (\$ MLN) Jul +1,161 +400.9 +310.4 - 54,703.0 PRODUCER PRICES (%) Jun - 0.2 +0.1 +0.2 119.6** - Yr / yr rise (%) +1.2 +1.5 +7.1 119.6 - INDUSTRIAL PROD . May - - - - - Yr / yr figures (%) - 3.2 +1.0 +9.8 108.4** - M4 MONEY SUPPLY (%) Jul +2.6 +4.2R +10.8 - - Total M4 adj . (trln pts) - - - 75.912 TRADE BALANCE Exports (bln pts) Jun 1,100.7 1,164.1 988.2 - - Imports (bln pts) May 1,315.7 1,433.4 1,236.5 - - Deficit / surplus May - 215.0 - 269.3 - 248.3 - - Deficit yr to date - 1,334.0 - 1,119.0 - 1,420.9 - - GOVT.BUDGET (bln pts) Govt.Fcast 96 Deficit / surplus Jul +282.1 - 380.6 +230.4 - Def . / surplus to date - 1,184.0 - 1,466.1 - 1,456.7 - 2.6 trln INDICATORS - quarterly QUARTER PVS QTR YR-AGO - - EPA Q2 +168,130 +31,230 +167,330 12.3 million GDP Yr-yr (%) Q1 +1.9 +2.3R +3.4 - - Absolute amount (trln pts) 18.1 17.8 16.9 - 69.7 INTEREST RATES Latest rate Pvs rate Date changed Key rate (%) 7.25 7.50 04/06/96 NOTES - [2 2 Bank of Spain 2] announces balance of payments . Jobless figures are registered unemployed at labour ministry . Trade data are customs-cleared , published by economy ministry . Latest M4 , currency reserves , industrial production data are provisional . GDP figures are quarterly on annualised basis . EPA - Quarterly survey of employment levels (INE) . Data give variation in employed persons , in thousands . Last column - TOTAL- is latest for jobless and accumulated for the rest (GDP total amount corresponds to 1995) . Government budget figures relate to central government finances only .

```
**General Consumer Price Index ( 100=1992 ) , Producer Prices  
Index and [3 3 Industrial Production Index 3] ( 100=1990 ) .  
1: 26667 Spain  
2: 604654 Bank_of_Spain  
3: 4199619 Industrial_production_index  
1: 26667 Spain  
    [pem_pos=2, final_scores=0.4973303,  
final_local_scores=0.6387479, pem=-0.18857051,  
similarity=24.176699]  
2: 604654 Bank_of_Spain  
    [pem_pos=0, final_scores=0.4474578,  
final_local_scores=0.565679, pem=0.0, similarity=20.970703]  
3: 4199619 Industrial_production_index  
    [pem_pos=0, final_scores=0.42756432,  
final_local_scores=0.62599474, pem=0.0, similarity=22.770645]
```

Sample: b'934_PRESS'

PRESS DIGEST - [1 1 Spain 1] - Aug 29 . Headlines from major national newspapers . [2 2 Reuters 2] has not verified these stories and does not vouch for their accuracy . [3 3 EL PAIS 3] - Judge accuses government of obstructing investigation into Lasa-Zabala (two of [4 4 GAL 4] victims) case [5 5 EL M UNDO 5] - Government wants to charge for prescriptions and some medical services [6 6 DIARIO 16 6] - Judge Javier Gomez de Liano says government is obstructing justice [7 7 ABC 7] - Prime Minister [8 8 Jose Maria Aznar 8] , positive assessment CINCO DIAS - BCH in the hive of [9 9 Chilean 9] pensions [10 1 0 EXPANSION 10] - Coopers and Lybrand emigrates to [11 11 Bas que Country 11] for fiscal reasons GACETA DE LOS NEGOCIOS - Government and [12 12 Catalan 12] nationalists set the scene for budget negotiations

```
1: 26667 Spain  
2: 18998750 Reuters  
3: 444409 El_País  
4: 539739 Grupos_Antiterroristas_de_Liberación  
5: 539568 El_Mundo_(Spain)  
6: 6242592 Diario_16  
7: 401572 ABC_(newspaper)  
8: 197618 José_María_Aznar  
9: 5489 Chile  
10: 6517516 Expansión  
11: 683036 Basque_Country_(greater_region)  
12: 6822 Catalonia  
1: 26667 Spain  
    [pem_pos=0, final_scores=0.6090372,  
final_local_scores=0.6199887, pem=-0.18857051,  
similarity=23.616888]  
2: 18998750 Reuters  
    [pem_pos=0, final_scores=0.31153035,  
final_local_scores=0.3642885, pem=0.0, similarity=14.960809]  
3: 444409 El_País  
    [pem_pos=0, final_scores=0.4127336,
```

```
final_local_scores=0.40262383, pem=0.0, similarity=16.10481]
4: 539739 Grupos_Antiterroristas_de_Liberación
    [pem_pos=0, final_scores=0.47873056,
final_local_scores=0.56005985, pem=-1.288439,
similarity=27.809748]
5: 539568 El_Mundo_(Spain)
    [pem_pos=0, final_scores=0.4897346,
final_local_scores=0.5033109, pem=0.0, similarity=19.109514]
6: 6242592 Diario_16
    [pem_pos=0, final_scores=0.47817045,
final_local_scores=0.5911611, pem=0.0, similarity=21.731138]
7: 401572 ABC_(newspaper)
    [pem_pos=11, final_scores=0.4081958,
final_local_scores=0.4113698, pem=-4.2527895,
similarity=39.493137]
8: 197618 José_María_Aznar
    [pem_pos=0, final_scores=0.5031208,
final_local_scores=0.53956497, pem=0.0, similarity=20.191408]
9: 5489 Chile
    [pem_pos=0, final_scores=0.55502117,
final_local_scores=0.5829821, pem=-0.25753683,
similarity=22.887587]
10: 6517516 Expansión
    [pem_pos=3, final_scores=0.48842335,
final_local_scores=0.6559047, pem=-1.609438,
similarity=32.415592]
11: 683036 Basque_Country_(greater_region)
    [pem_pos=1, final_scores=0.17109214,
final_local_scores=0.16861129, pem=-0.72269726,
similarity=13.05155]
12: 6822 Catalonia
    [pem_pos=1, final_scores=0.405245,
final_local_scores=0.39374906, pem=-1.5206007,
similarity=24.109232]
```

Sample: b'935_Lenzing'

Lenzing expects negative results in H2 . [1 1 VIENNA 1]
1996-08-29 [2 2 Austrian 2] viscose fibre maker [3 3 Lenzing A
G 3] said on Thursday it expected to post negative group
results in the second half of the year after posting losses
in the first six months . " A preview of the second half of
1996 does not reveal any signs of a significant improvement
in market conditions , " Lenzing said in a statement released
ahead of its earnings conference . For the first six months
of the year , Lenzing said it posted a group pre-tax loss of
84.5 million schillings from a profit of 160 million in the
year-ago period . The group attributed the first-half losses
to weak demand and falling prices of viscose fibres , as well
as sluggish economies in the West . -- Julia Ferguson , [4 4 V
ienna 4] newsroom , +431 53112 274

1: 55866 Vienna

2: 26964606 Austria

```
3: 976292 Lenzing_AG
4: 55866 Vienna
1: 55866 Vienna
    [pem_pos=0, final_scores=0.5856183,
final_local_scores=0.5575515, pem=-0.12517071,
similarity=21.40886]
2: 26964606 Austria
    [pem_pos=0, final_scores=0.683335,
final_local_scores=0.67536217, pem=-0.40674555,
similarity=26.455812]
3: 976292 Lenzing_AG
    [pem_pos=0, final_scores=0.36023706,
final_local_scores=0.5062341, pem=0.0, similarity=19.196749]
4: 55866 Vienna
    [pem_pos=0, final_scores=0.5456423,
final_local_scores=0.5138838, pem=-0.12517071,
similarity=20.105728]
```

Sample: b'936_Algeria'

[1 1 Algeria 1] faults [2 2 Britain 2] over [3 3 Islamists 3] gathering . [4 4 PARIS 4] 1996-08-28 [5 5 Algeria 5] , fighting a vicious war against [6 6 Moslem 6] fundamentalist guerrillas , attacked [7 7 Britain 7] on Wednesday for allowing [8 8 Islamist 8] groups to meet in [9 9 London 9] . The [10 10 Islamist 10] gathering , due to be held in [11 11 London 11] on September 8 , has triggered concern and anger in several other [12 12 Arab 12] countries like [13 13 Egypt 13] which is also fighting armed [14 14 Moslem 14] fundamentalists . [1 1 British 1] [2 Jewish 2] groups have also voiced protest because they said Palestinian Islamist [15 15 Hamas 15] as well as the banned [16 16 Algerian 16] [17 17 Islamic Salvation Front 17] ([18 18 FIS 18]) are among those radical [19 19 Islamists 19] attending the conference . A foreign ministry spokesman said in a statement read on [20 20 Algerian 20] television that [21 21 Algeria 21] " has received with concern the information over a meeting of terrorist groups working against the interests of the [22 22 Arab 22] and [23 23 Islamic 23] world . " " [24 24 Algeria 24] expresses its sharp rejection of a meeting putting together masterminds and ideologists and financers of terrorism , " the spokesman said , adding the [25 25 Algerian 25] government has asked the [26 26 British 26] embassy in [27 27 Algiers 27] for clarifications . The [28 28 Algerian 28] ambassador in [29 29 London 29] has also asked for clarification from the Foreign Office over the meeting of [30 30 Islamist 30] groups . [31 31 Algeria 31] said " they are clearly working to undermine the stability " of [32 32 Arab 32] countries . [33 33 British 33] Foreign Secretary [34 34 Malcolm Rifkind 34] said on Tuesday from [35 35 Pakistan 35] his government would only take action against the planned [36 36 Islamists 36] gathering in [37 37 London 37] if [38 38 British 38] law was broken . " People who wish to hold conferences of

course do n't need to seek permission from the government in [39 39 Britain 39] , " he said . An estimated 50,000 people , including more than 110 foreigners , have been killed in [40 40 Algeria 40] 's violence pitting [41 41 Moslem 41] rebels against government forces since early 1992 when authorities in [42 42 Algeria 42] cancelled a general election in which [43 43 FIS 43] had taken a commanding lead .

1: 358 Algeria
2: 31717 United_Kingdom
3: 15012 Islamism
4: 22989 Paris
5: 358 Algeria
6: 6037917 Islam
7: 31717 United_Kingdom
8: 15012 Islamism
9: 17867 London
10: 15012 Islamism
11: 17867 London
12: 159433 Arab_world
13: 8087628 Egypt
14: 6037917 Islam
15: 13913 Hamas
16: 358 Algeria
17: 403399 Islamic_Salvation_Front
18: 403399 Islamic_Salvation_Front
19: 15012 Islamism
20: 358 Algeria
21: 358 Algeria
22: 159433 Arab_world
23: 6037917 Islam
24: 358 Algeria
25: 358 Algeria
26: 31717 United_Kingdom
27: 358 Algeria
28: 358 Algeria
29: 17867 London
30: 15012 Islamism
31: 358 Algeria
32: 159433 Arab_world
33: 31717 United_Kingdom
34: 354551 Malcolm_Rifkind
35: 23235 Pakistan
36: 15012 Islamism
37: 17867 London
38: 31717 United_Kingdom
39: 31717 United_Kingdom
40: 358 Algeria
41: 6037917 Islam
42: 358 Algeria
43: 403399 Islamic_Salvation_Front
1: 5380456 British_Jews
[winner: 31717 United_Kingdom, pem_pos=0,

```
final_scores=0.25515038, final_local_scores=0.21201718,
pem=-0.48936236, similarity=13.077957]
    RECALL MISS: 5380456 British_Jews]
        [winner: 25955086 Jews, pem_pos=1,
final_scores=0.1643139, final_local_scores=0.14890224,
pem=-0.80476934, similarity=12.909714]
        gt: 5380456 British_Jews, pem_pos=8,
final_scores=-0.6100367, final_local_scores=-0.64364076,
pem=-3.7875285, similarity=5.479359 ]
1: 358 Algeria
    [pem_pos=0, final_scores=0.5527974,
final_local_scores=0.53097016, pem=-0.33694586,
similarity=21.767286]
2: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5904125,
final_local_scores=0.5782401, pem=-0.61114657,
similarity=24.66906]
3: 15012 Islamism
    [pem_pos=0, final_scores=0.33689892,
final_local_scores=0.32402754, pem=0.0, similarity=13.759341]
4: 22989 Paris
    [pem_pos=0, final_scores=0.48771518,
final_local_scores=0.51241463, pem=-0.115301244,
similarity=20.008213]
5: 358 Algeria
    [pem_pos=0, final_scores=0.6504524,
final_local_scores=0.6376434, pem=-0.33694586,
similarity=24.950628]
6: 6037917 Islam
    [pem_pos=1, final_scores=0.46735024,
final_local_scores=0.4758463, pem=-1.2150909,
similarity=24.897768]
7: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5330121,
final_local_scores=0.5155389, pem=-0.61114657,
similarity=22.797932]
8: 15012 Islamism
    [pem_pos=0, final_scores=0.33769435,
final_local_scores=0.3248964, pem=0.0, similarity=13.78527]
9: 17867 London
    [pem_pos=0, final_scores=0.5882227,
final_local_scores=0.6031134, pem=-0.09086737,
similarity=22.58197]
10: 15012 Islamism
    [pem_pos=0, final_scores=0.3343367,
final_local_scores=0.32122868, pem=0.0, similarity=13.6758175]
11: 17867 London
    [pem_pos=0, final_scores=0.5275124,
final_local_scores=0.5367966, pem=-0.09086737,
similarity=20.602945]
12: 159433 Arab_world
    [pem_pos=14, final_scores=0.48516196,
```

```
final_local_scores=0.4615872, pem=-3.8052368,
similarity=38.557865]
13: 8087628 Egypt
    [pem_pos=0, final_scores=0.53950006,
final_local_scores=0.50529426, pem=-0.27480945,
similarity=20.663158]
14: 6037917 Islam
    [pem_pos=1, final_scores=0.37150186,
final_local_scores=0.3711465, pem=-1.2150909,
similarity=21.773317]
15: 13913 Hamas
    [pem_pos=0, final_scores=0.45515847,
final_local_scores=0.48468417, pem=0.0, similarity=18.553656]
16: 358 Algeria
    [pem_pos=0, final_scores=0.64401716,
final_local_scores=0.63061386, pem=-0.20917186,
similarity=24.045998]
17: 403399 Islamic_Salvation_Front
    [pem_pos=0, final_scores=0.54743624,
final_local_scores=0.6109096, pem=0.0, similarity=22.320473]
18: 403399 Islamic_Salvation_Front
    [pem_pos=2, final_scores=0.57521117,
final_local_scores=0.6412495, pem=-1.4075395,
similarity=30.880295]
19: 15012 Islamism
    [pem_pos=0, final_scores=0.264893,
final_local_scores=0.24537197, pem=0.0, similarity=11.412102]
20: 358 Algeria
    [pem_pos=0, final_scores=0.66007406,
final_local_scores=0.6481536, pem=-0.20917186,
similarity=24.569418]
21: 358 Algeria
    [pem_pos=0, final_scores=0.6770217,
final_local_scores=0.66666627, pem=-0.33694586,
similarity=25.816729]
22: 159433 Arab_world
    [pem_pos=14, final_scores=0.55814224,
final_local_scores=0.5413071, pem=-3.8052368,
similarity=40.936867]
23: 6037917 Islam
    [pem_pos=0, final_scores=0.3800661,
final_local_scores=0.38050163, pem=-0.21141872,
similarity=16.594368]
24: 358 Algeria
    [pem_pos=0, final_scores=0.64098877,
final_local_scores=0.6273058, pem=-0.33694586,
similarity=24.642134]
25: 358 Algeria
    [pem_pos=0, final_scores=0.76121503,
final_local_scores=0.7586347, pem=-0.20917186,
similarity=27.866394]
26: 31717 United_Kingdom
```

```
[pem_pos=0, final_scores=0.5594494,
final_local_scores=0.54441756, pem=-0.48936236,
similarity=22.997448]
27: 358 Algeria
    [pem_pos=8, final_scores=0.26962113,
final_local_scores=0.22259328, pem=-2.7421203,
similarity=25.644417]
28: 358 Algeria
    [pem_pos=0, final_scores=0.757206,
final_local_scores=0.7542555, pem=-0.20917186,
similarity=27.735708]
29: 17867 London
    [pem_pos=0, final_scores=0.5557587,
final_local_scores=0.56765145, pem=-0.09086737,
similarity=21.523714]
30: 15012 Islamism
    [pem_pos=0, final_scores=0.43683368,
final_local_scores=0.43319106, pem=0.0, similarity=17.016998]
31: 358 Algeria
    [pem_pos=0, final_scores=0.6284003,
final_local_scores=0.61355484, pem=-0.33694586,
similarity=24.231777]
32: 159433 Arab_world
    [pem_pos=14, final_scores=0.56909984,
final_local_scores=0.55327654, pem=-3.8052368,
similarity=41.29406]
33: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5596555,
final_local_scores=0.54464275, pem=-0.48936236,
similarity=23.004168]
34: 354551 Malcolm_Rifkind
    [pem_pos=0, final_scores=0.57709116,
final_local_scores=0.69645184, pem=0.0, similarity=24.873224]
35: 23235 Pakistan
    [pem_pos=0, final_scores=0.5874423,
final_local_scores=0.58927506, pem=-0.21944217,
similarity=22.868216]
36: 15012 Islamism
    [pem_pos=0, final_scores=0.35293782,
final_local_scores=0.3415476, pem=0.0, similarity=14.282175]
37: 17867 London
    [pem_pos=0, final_scores=0.5565191,
final_local_scores=0.56848204, pem=-0.09086737,
similarity=21.548502]
38: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.54744303,
final_local_scores=0.5313025, pem=-0.48936236,
similarity=22.606068]
39: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.54485315,
final_local_scores=0.52847344, pem=-0.61114657,
similarity=23.183926]
```

```
40: 358 Algeria
    [pem_pos=0, final_scores=0.6350417,
final_local_scores=0.62080956, pem=-0.33694586,
similarity=24.448273]
41: 6037917 Islam
    [pem_pos=1, final_scores=0.44510323,
final_local_scores=0.45154482, pem=-1.2150909,
similarity=24.172564]
42: 358 Algeria
    [pem_pos=0, final_scores=0.6180411,
final_local_scores=0.60223895, pem=-0.33694586,
similarity=23.894089]
43: 403399 Islamic_Salvation_Front
    [pem_pos=2, final_scores=0.50416595,
final_local_scores=0.5636434, pem=-1.4075395,
similarity=28.564375]
1: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.25515038,
final_local_scores=0.21201718, pem=-0.48936236,
similarity=13.077957]
2: 25955086 Jews
    [pem_pos=1, final_scores=0.1643139,
final_local_scores=0.14890224, pem=-0.80476934,
similarity=12.909714]
```

Sample: b'937_OFFICIAL'

OFFICIAL JOURNAL CONTENTS - OJ L 218 OF AUGUST 28 , 1996 . *

Commission Regulation (EC) No 1676/96 of 30 July 1996
amending Regulation ([1 1 EEC 1]) No 2454/93 laying down
provisions for the implementation of Council Regulation ([2 2
EEC 2]) No 2913/92 establishing the Community Customs Code
END OF DOCUMENT .

```
1: 9578 European_Economic_Community
2: 9578 European_Economic_Community
1: 9578 European_Economic_Community
    [pem_pos=0, final_scores=0.455476,
final_local_scores=0.3820567, pem=0.0, similarity=15.491047]
2: 9578 European_Economic_Community
    [pem_pos=0, final_scores=0.4761967,
final_local_scores=0.39208108, pem=0.0, similarity=15.7901945]
```

Sample: b'938_Wall'

[1 1 Wall Street 1] ponders [2 2 Rubin 2] 's role if [3 3 Clinton 3] wins . Donna Sells [4 4 NEW YORK 4] The outcome of the November elections emerged as a hot topic on [5 5 Wall Street 5] this week as financial pundits debated whether [6 6 Robert Rubin 6] might forgo a second term as [7 7 Treasury 7] secretary if President [8 8 Clinton 8] is re-elected . Concern centred on the currency markets since [9 9 Rubin 9] 's tour de force has been his unflagging support of the dollar . Speculation that [10 10 Rubin 10] might not stay in his post grew after he sidestepped questions about any future

[11 11 Cabinet 11] post during television interviews at the [1 2 12 Democratic 12] convention in [13 13 Chicago 13] this week . Should [14 14 Rubin 14] leave , [15 15 Wall Street 15] would worry that he might take his strong-dollar policy with him . [16 16 Rubin 16] 's predecessor at the [17 17 Treasury 1 7] , [18 18 Lloyd Bentsen 18] , was viewed with suspicion by some in the financial markets who thought he had tried to push down the dollar to gain an edge in trade negotiations with [19 19 Japan 19] . " Obviously , under the [20 20 Clinton 20] administration , we 've seen two distinctively different dollar policies , " said Chris Widness , an international economist at Chase Securities Inc. " Under [21 21 Rubin 21] , the [22 22 U.S. 22] has certainly looked for a strong dollar . " That strategy , backed up by timely instances of joint central bank intervention , helped the dollar battle back from post-Second World War lows of 1.3438 [23 23 German 2 3] marks on March 8 , 1995 , and 79.75 [24 24 Japanese 24] yen on April 19 , 1995 . Currently , the dollar stands at about 1.48 marks and 109 yen . [25 25 Rubin 25] was widely hailed as the architect of the dollar 's comeback , using skills and expertise gained in 26 years on [26 26 Wall Street 26] , part of which were spent as co-chairman of Goldman , Sachs and Co . Inc . " [27 27 Rubin 27] has done a fine job in that position , " said Michael Faust , a portfolio manager at Bailard , Biehl and Kaiser , which manages just under \$ 1 billion in global stocks and bonds . " Anyone who would come in there to replace him would have awfully big shoes to fill . " Fear that a new [28 28 Treasury 28] secretary might favour a return to Bentsen-era policy could spell trouble for financial markets . Some overseas investors might shy away from buying [29 29 U.S. 29] stocks and bonds or even sell them when the dollar is weakening . As for [30 30 U.S. Treasury 30] securities , Widness explained that [31 31 Alan Greenspan 31] 's reappointment as chairman of the [32 32 Federal Reserve 32] and the outlook for the federal budget were more important than whether [33 33 Rubin 33] continues at the [34 3 4 Treasury 34] . " Although , if we did get someone that was seen as looking for a dollar depreciation , it would probably hurt capital flows to the [35 35 United States 35] , " said Widness , adding that could hurt [36 36 U.S. 36] stocks and , to a lesser degree , bonds . Still , markets may have little to fear from any [37 37 Rubin 37] successor because the firm dollar policy has yielded positive results . If that is true , then any new [38 38 Treasury 38] chief would need to be as effective as [39 39 Rubin 39] in convincing markets that the [40 40 White House 40] does indeed want a strong currency . " If he left , the first question people would ask the next guy is , ' What 's your view on the dollar ? ' " said Michael Perelstein , portfolio manager of MainStay International Funds . " And all I can say as a piece of advice is that they 'd better say exactly the same thing (as [41 41 Rubin 41]) , if not stronger , " Perelstein said . "

Otherwise , you get selling out of [42 42 Tokyo 42] and [43 43 Frankfurt 43] again . "

1: 37274 Wall_Street
2: 614418 Robert_Rubin
3: 3356 Bill_Clinton
4: 645042 New_York_City
5: 37274 Wall_Street
6: 614418 Robert_Rubin
7: 53667 United_States_Department_of_the_Treasury
8: 3356 Bill_Clinton
9: 614418 Robert_Rubin
10: 614418 Robert_Rubin
11: 32187890 Cabinet_(government)
12: 5043544 Democratic_Party_(United_States)
13: 6886 Chicago
14: 614418 Robert_Rubin
15: 37274 Wall_Street
16: 614418 Robert_Rubin
17: 53667 United_States_Department_of_the_Treasury
18: 299257 Lloyd_Bentsen
19: 15573 Japan
20: 3356 Bill_Clinton
21: 614418 Robert_Rubin
22: 3434750 United_States
23: 11867 Germany
24: 15573 Japan
25: 614418 Robert_Rubin
26: 37274 Wall_Street
27: 614418 Robert_Rubin
28: 53667 United_States_Department_of_the_Treasury
29: 3434750 United_States
30: 53667 United_States_Department_of_the_Treasury
31: 161947 Alan_Greenspan
32: 10819 Federal_Reserve_System
33: 614418 Robert_Rubin
34: 53667 United_States_Department_of_the_Treasury
35: 3434750 United_States
36: 3434750 United_States
37: 614418 Robert_Rubin
38: 53667 United_States_Department_of_the_Treasury
39: 614418 Robert_Rubin
40: 33057 White_House
41: 614418 Robert_Rubin
42: 30057 Tokyo
43: 10992 Frankfurt

1: 37274 Wall_Street
[pem_pos=0, final_scores=0.42714006,
final_local_scores=0.4103107, pem=-0.039525714,
similarity=16.54915]
2: 614418 Robert_Rubin
[pem_pos=0, final_scores=0.92169267,
final_local_scores=0.912298, pem=0.0, similarity=31.314507]

```
3: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.82107097,
final_local_scores=0.84628695, pem=0.0, similarity=29.344604]
4: 645042 New_York_City
    [pem_pos=1, final_scores=0.3532833,
final_local_scores=0.36274529, pem=-1.7906687,
similarity=24.65269]
5: 37274 Wall_Street
    [pem_pos=0, final_scores=0.34288305,
final_local_scores=0.31827268, pem=-0.039525714,
similarity=13.802551]
6: 614418 Robert_Rubin
    [pem_pos=0, final_scores=0.7631927,
final_local_scores=0.73777694, pem=0.0, similarity=26.10645]
7: 53667 United_States_Department_of_the_Treasury
    [pem_pos=0, final_scores=0.4693594,
final_local_scores=0.45617998, pem=-1.0120404,
similarity=23.206667]
8: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.51611906,
final_local_scores=0.510166, pem=-1.920079,
similarity=29.755774]
9: 614418 Robert_Rubin
    [pem_pos=0, final_scores=0.86252993,
final_local_scores=0.8476716, pem=0.0, similarity=29.385925]
10: 614418 Robert_Rubin
    [pem_pos=0, final_scores=1.0036198,
final_local_scores=1.001791, pem=0.0, similarity=33.985153]
11: 32187890 Cabinet_(government)
    [pem_pos=1, final_scores=0.33552098,
final_local_scores=0.39104408, pem=-2.1059742,
similarity=27.211863]
12: 5043544 Democratic_Party_(United_States)
    [pem_pos=0, final_scores=0.4209358,
final_local_scores=0.448196, pem=-0.076093495,
similarity=17.878584]
13: 6886 Chicago
    [pem_pos=0, final_scores=0.61673397,
final_local_scores=0.66115993, pem=-0.20658566,
similarity=24.94349]
14: 614418 Robert_Rubin
    [pem_pos=0, final_scores=0.8972319,
final_local_scores=0.88557833, pem=0.0, similarity=30.517136]
15: 37274 Wall_Street
    [pem_pos=0, final_scores=0.39309174,
final_local_scores=0.37311804, pem=-0.039525714,
similarity=15.439247]
16: 614418 Robert_Rubin
    [pem_pos=0, final_scores=0.9367458,
final_local_scores=0.9287413, pem=0.0, similarity=31.805206]
17: 53667 United_States_Department_of_the_Treasury
    [pem_pos=0, final_scores=0.4356926,
```

```
final_local_scores=0.4194041, pem=-1.0120404,
similarity=22.1092]
18: 299257 Lloyd_Bentsen
    [pem_pos=0, final_scores=0.55906564,
final_local_scores=0.6022673, pem=0.0, similarity=22.06257]
19: 15573 Japan
    [pem_pos=0, final_scores=0.5276707,
final_local_scores=0.6585762, pem=-0.13549899,
similarity=24.479803]
20: 3356 Bill_Clinton
    [pem_pos=0, final_scores=0.83731776,
final_local_scores=0.8640341, pem=0.0, similarity=29.874214]
21: 614418 Robert_Rubin
    [pem_pos=0, final_scores=0.9334559,
final_local_scores=0.9251475, pem=0.0, similarity=31.69796]
22: 3434750 United_States
    [pem_pos=0, final_scores=0.52252454,
final_local_scores=0.56337434, pem=-0.14075686,
similarity=21.667385]
23: 11867 Germany
    [pem_pos=1, final_scores=0.5939093,
final_local_scores=0.70190257, pem=-1.0465823,
similarity=30.727365]
24: 15573 Japan
    [pem_pos=0, final_scores=0.63981694,
final_local_scores=0.7810789, pem=-0.99076414,
similarity=32.786594]
25: 614418 Robert_Rubin
    [pem_pos=0, final_scores=0.87659854,
final_local_scores=0.86303943, pem=0.0, similarity=29.844532]
26: 37274 Wall_Street
    [pem_pos=0, final_scores=0.42245513,
final_local_scores=0.40519315, pem=-0.039525714,
similarity=16.396433]
27: 614418 Robert_Rubin
    [pem_pos=0, final_scores=0.945231,
final_local_scores=0.93801004, pem=0.0, similarity=32.081802]
28: 53667 United_States_Department_of_the_Treasury
    [pem_pos=0, final_scores=0.3827023,
final_local_scores=0.36152023, pem=-1.0120404,
similarity=20.38183]
29: 3434750 United_States
    [pem_pos=0, final_scores=0.45256948,
final_local_scores=0.48695904, pem=-0.14075686,
similarity=19.387]
30: 53667 United_States_Department_of_the_Treasury
    [pem_pos=0, final_scores=0.44643354,
final_local_scores=0.43113697, pem=0.0, similarity=16.9557]
31: 161947 Alan_Greenspan
    [pem_pos=0, final_scores=0.4810769,
final_local_scores=0.48461974, pem=0.0, similarity=18.551733]
32: 10819 Federal_Reserve_System
```

```
[pem_pos=0, final_scores=0.33250004,
final_local_scores=0.31420416, pem=0.0, similarity=13.466191]
33: 614418 Robert_Rubin
    [pem_pos=0, final_scores=0.90283793,
final_local_scores=0.8917021, pem=0.0, similarity=30.699884]
34: 53667 United_States_Department_of_the_Treasury
    [pem_pos=0, final_scores=0.45144194,
final_local_scores=0.4366079, pem=-1.0120404,
similarity=22.622597]
35: 3434750 United_States
    [pem_pos=0, final_scores=0.47251844,
final_local_scores=0.50875026, pem=-0.073457584,
similarity=19.67131]
36: 3434750 United_States
    [pem_pos=0, final_scores=0.47738296,
final_local_scores=0.514064, pem=-0.14075686,
similarity=20.195868]
37: 614418 Robert_Rubin
    [pem_pos=0, final_scores=0.9258242,
final_local_scores=0.91681105, pem=0.0, similarity=31.449184]
38: 53667 United_States_Department_of_the_Treasury
    [pem_pos=0, final_scores=0.4477771,
final_local_scores=0.43260455, pem=-1.0120404,
similarity=22.503128]
39: 614418 Robert_Rubin
    [pem_pos=0, final_scores=0.99089104,
final_local_scores=0.9878867, pem=0.0, similarity=33.570225]
40: 33057 White_House
    [pem_pos=0, final_scores=0.49721545,
final_local_scores=0.49835336, pem=-0.0056634876,
similarity=18.99237]
41: 614418 Robert_Rubin
    [pem_pos=0, final_scores=0.8945027,
final_local_scores=0.882597, pem=0.0, similarity=30.42817]
42: 30057 Tokyo
    [pem_pos=0, final_scores=0.4272912,
final_local_scores=0.51784694, pem=-0.09283106,
similarity=20.048128]
43: 10992 Frankfurt
    [pem_pos=0, final_scores=0.2967829,
final_local_scores=0.35675666, pem=-0.07062824,
similarity=15.120131]
```

Sample: b'939_SOCCER'

```
SOCCER - [1 1 AUSTRIA 1] BEAT [2 2 SCOTLAND 2] 4-0 IN [3 3 EU
ROPEAN 3] UNDER-21 MATCH . [4 4 AMSTETTEN 4] , [5 5 Austria 5]
1996-08-30 [1 1 Austria 1] beat [2 2 Scotland 2] 4-0
( halftime 3-0 ) in a [6 6 European 6] under-21 championship
match on Friday . Scorers : [7 7 Ewald Brenner 7] ( 5th
minute ) , Mario Stieglmaier ( 42nd ) , [8 8 Ronald Brunmayr 8]
( 43rd and 56th ) . Attendance : 800
1: 695226 Austria_national_football_team
```

```
2: 578101 Scotland_national_football_team
3: 9239 Europe
4: 4745273 Amstetten,_Lower_Austria
5: 26964606 Austria
6: 9239 Europe
7: 24425438 Ewald_Brenner
8: 19220665 Ronald_Brunmayr
1: 21478893 Austria_national_under-21_football_team
    [winner: 695226 Austria_national_football_team,
pem_pos=1, final_scores=0.061219558,
final_local_scores=-0.0061021447, pem=-2.931263,
similarity=19.848276]
    RECALL MISS: 21478893
Austria_national_under-21_football_team]
2: 4683429 Scotland_national_under-21_football_team
    [winner: 578101 Scotland_national_football_team,
pem_pos=1, final_scores=0.23357202,
final_local_scores=0.20089424, pem=-2.6956348,
similarity=24.74408]
    gt: 4683429 Scotland_national_under-21_football_team,
pem_pos=17, final_scores=0.14192061,
final_local_scores=0.11994678, pem=-4.548483,
similarity=32.404522 ]
1: 695226 Austria_national_football_team
    [pem_pos=1, final_scores=0.47084707,
final_local_scores=0.45333892, pem=-2.931263,
similarity=33.558914]
2: 578101 Scotland_national_football_team
    [pem_pos=1, final_scores=0.42919284,
final_local_scores=0.39497244, pem=-2.6956348,
similarity=30.535759]
3: 9239 Europe
    [pem_pos=0, final_scores=0.32840747,
final_local_scores=0.38134003, pem=-0.25557294,
similarity=16.859505]
4: 4745273 Amstetten,_Lower_Austria
    [pem_pos=0, final_scores=0.4068299,
final_local_scores=0.517375, pem=-0.27663225,
similarity=21.033585]
5: 26964606 Austria
    [pem_pos=0, final_scores=0.5407486,
final_local_scores=0.57339734, pem=-0.25262418,
similarity=22.574842]
6: 9239 Europe
    [pem_pos=0, final_scores=0.50550455,
final_local_scores=0.57479167, pem=-0.25557294,
similarity=22.632488]
7: 24425438 Ewald_Brenner
    [pem_pos=0, final_scores=0.49030983,
final_local_scores=0.62993914, pem=0.0, similarity=22.888353]
8: 19220665 Ronald_Brunmayr
    [pem_pos=0, final_scores=0.60129195,
```

```
final_local_scores=0.70468295, pem=0.0, similarity=25.118858]
1: 695226 Austria_national_football_team
    [pem_pos=1, final_scores=0.061219558,
final_local_scores=-0.0061021447, pem=-2.931263,
similarity=19.848276]
2: 578101 Scotland_national_football_team
    [pem_pos=1, final_scores=0.23357202,
final_local_scores=0.20089424, pem=-2.6956348,
similarity=24.74408]
```

Sample: b'940_SOCCER'

SOCCER - [1 1 WALES 1] BEAT [2 2 SAN MARINO 2] 4-0 IN
UNDER-21 MATCH . [3 3 BARRY 3] , [4 4 Wales 4] 1996-08-30 [1 1
Wales 1] beat [5 5 San Marino 5] 4-0 (halftime 2-0) in a [6
6 European 6] under-21 soccer match on Friday . Scorers : [2 2
Wales 2] - [7 7 John Hartson 7] (12th , 56th and 83rd
minutes) , [8 8 Scott Young 8] (24th) Attendance : 1,800

```
1: 580291 Wales_national_football_team
2: 654427 San_Marino_national_football_team
3: 1251024 Barry,_Vale_of_Glamorgan
4: 69894 Wales
5: 654427 San_Marino_national_football_team
6: 9239 Europe
7: 1062932 John_Hartson
8: 8908579 Scott_Young_(Welsh_footballer)
1: 5082938 Wales_national_under-21_football_team
    [winner: 580291 Wales_national_football_team, pem_pos=1,
final_scores=0.12148969, final_local_scores=0.03609076,
pem=-2.4526134, similarity=18.504425]
    gt: 5082938 Wales_national_under-21_football_team,
pem_pos=13, final_scores=0.063333645,
final_local_scores=-0.011852324, pem=-4.354265,
similarity=27.415188 ]
2: 5082938 Wales_national_under-21_football_team
    [winner: 580291 Wales_national_football_team, pem_pos=1,
final_scores=0.2796772, final_local_scores=0.2117267,
pem=-2.4526134, similarity=23.745752]
    gt: 5082938 Wales_national_under-21_football_team,
pem_pos=13, final_scores=0.16049497,
final_local_scores=0.1076808, pem=-4.354265,
similarity=30.982294 ]
1: 580291 Wales_national_football_team
    [pem_pos=1, final_scores=0.32777262,
final_local_scores=0.2614235, pem=-2.4526134,
similarity=25.228804]
2: 654427 San_Marino_national_football_team
    [pem_pos=1, final_scores=0.5064959,
final_local_scores=0.5201482, pem=-2.2175765,
similarity=31.6715]
3: 1251024 Barry,_Vale_of_Glamorgan
    [pem_pos=0, final_scores=0.38338178,
final_local_scores=0.45502657, pem=-2.1395404,
```

```
similarity=29.303766]
4: 69894 Wales
    [pem_pos=0, final_scores=0.5975734,
final_local_scores=0.6432191, pem=-0.37222588,
similarity=25.308876]
5: 654427 San_Marino_national_football_team
    [pem_pos=1, final_scores=0.41769916,
final_local_scores=0.42315125, pem=-2.2175765,
similarity=28.776917]
6: 9239 Europe
    [pem_pos=0, final_scores=0.56619954,
final_local_scores=0.6836332, pem=-0.25557294,
similarity=25.880537]
7: 1062932 John_Hartson
    [pem_pos=0, final_scores=0.8329999,
final_local_scores=0.9008228, pem=0.0, similarity=30.972061]
8: 8908579 Scott_Young_(Welsh_footballer)
    [pem_pos=1, final_scores=0.40867072,
final_local_scores=0.44962484, pem=-0.91140074,
similarity=22.463755]
1: 580291 Wales_national_football_team
    [pem_pos=1, final_scores=0.12148969,
final_local_scores=0.03609076, pem=-2.4526134,
similarity=18.504425]
2: 580291 Wales_national_football_team
    [pem_pos=1, final_scores=0.2796772,
final_local_scores=0.2117267, pem=-2.4526134,
similarity=23.745752]
```

Sample: b'941_CYCLING'
CYCLING - [1 1 BALLANGER 1] KEEPS SPRINT TITLE IN STYLE .
Martin Ayres [2 2 MANCHESTER 2] , [3 3 England 3] 1996-08-30 [4 4 Felicia Ballanger 4] of [5 5 France 5] confirmed her status as the world 's number one woman sprinter when she retained her title at the world cycling championships on Friday . [6 6 Ballanger 6] beat [7 7 Germany 7] 's [8 8 Annett Neumann 8] 2-0 in the best-of-three matches final to add the world title to the 1 Olympic gold medal she won in July . [9 9 France 9] also took third place in the sprint , Magali Faure defeating ex-world champion [10 10 Tanya Dubnicoff 10] of [11 11 Canada 11] 2-0 . [12 12 Ballanger 12] , 25 , will be aiming to complete a track double when she defends her 500 metres time trial title on Saturday . The other final of the night , the women 's 24-kms points race , also ended in success for the reigning champion . [13 13 Russia 13] 's Svetlana Samokhalova fought off a spirited challenge from [14 14 American 14] Jane Quigley to take the title for a second year . [15 15 Russia 15] , the only nation to have two riders in the field , made full use of their numerical superiority . Goulnara Fatkoullina helped Samokhalova to build an unbeatable points lead before snatching the bronze medal . Quigley , a former medallist in the points event , led the

race at half distance . " I went so close this time , but having two riders certainly gave the [16 16 Russians 16] an advantage , " she said . The first six riders lapped the field , which left former world champion Ingrid Haringa of the [17 17 Netherlands 17] down in seventh place despite having the second highest points score . 2 Olympic champion [1 8 18 Nathalie Lancien 18] of [19 19 France 19] also missed the winning attack and finished a disappointing 10th .

1: 6519486 Félicia_Ballanger
2: 20206 Manchester
3: 9316 England
4: 6519486 Félicia_Ballanger
5: 5843419 France
6: 6519486 Félicia_Ballanger
7: 11867 Germany
8: 6691583 Annett_Neumann
9: 5843419 France
10: 11451462 Tanya_Dubnicoff
11: 5042916 Canada
12: 6519486 Félicia_Ballanger
13: 25391 Russia
14: 3434750 United_States
15: 25391 Russia
16: 25391 Russia
17: 21148 Netherlands
18: 17825936 Nathalie_Lancien
19: 5843419 France
1: 72311 1996_Summer_Olympics
[winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.21577004, final_local_scores=-0.25926843,
pem=-3.2532303, similarity=14.044199]
RECALL MISS: 72311 1996_Summer_Olympics]
2: 72311 1996_Summer_Olympics
[winner: 27566 Summer_Olympic_Games, pem_pos=2,
final_scores=-0.13318326, final_local_scores=-0.16905493,
pem=-3.2532303, similarity=16.73635]
RECALL MISS: 72311 1996_Summer_Olympics]
1: 6519486 Félicia_Ballanger
[pem_pos=0, final_scores=0.47687626,
final_local_scores=0.572182, pem=0.0, similarity=21.164764]
2: 20206 Manchester
[pem_pos=0, final_scores=0.4924826,
final_local_scores=0.5188075, pem=-0.3209542,
similarity=21.317362]
3: 9316 England
[pem_pos=0, final_scores=0.65186834,
final_local_scores=0.67645997, pem=-0.19053856,
similarity=25.312805]
4: 6519486 Félicia_Ballanger
[pem_pos=0, final_scores=0.711109,
final_local_scores=0.82804567, pem=0.0, similarity=28.800247]
5: 5843419 France

```
[pem_pos=0, final_scores=0.77600855,
final_local_scores=0.7602381, pem=-0.16240834,
similarity=27.659935]
6: 6519486 Félicia_Ballanger
[pem_pos=0, final_scores=0.6130836,
final_local_scores=0.7209678, pem=0.0, similarity=25.604832]
7: 11867 Germany
[pem_pos=0, final_scores=0.863954,
final_local_scores=0.8813917, pem=-0.15529114,
similarity=31.236698]
8: 6691583 Annett_Neumann
[pem_pos=0, final_scores=0.5668935,
final_local_scores=0.740125, pem=0.0, similarity=26.17652]
9: 5843419 France
[pem_pos=0, final_scores=0.75213045,
final_local_scores=0.7341549, pem=-0.16240834,
similarity=26.88156]
10: 11451462 Tanya_Dubnicoff
[pem_pos=0, final_scores=0.6247349,
final_local_scores=0.77721345, pem=0.0, similarity=27.283312]
11: 5042916 Canada
[pem_pos=0, final_scores=0.8166531,
final_local_scores=0.8340378, pem=-0.11662793,
similarity=29.613304]
12: 6519486 Félicia_Ballanger
[pem_pos=0, final_scores=0.6387885,
final_local_scores=0.7490465, pem=0.0, similarity=26.442755]
13: 25391 Russia
[pem_pos=0, final_scores=0.8224538,
final_local_scores=0.85926753, pem=-0.20299491,
similarity=30.835888]
14: 3434750 United_States
[pem_pos=0, final_scores=0.6922437,
final_local_scores=0.7071182, pem=-0.12994486,
similarity=25.898191]
15: 25391 Russia
[pem_pos=0, final_scores=0.48101097,
final_local_scores=0.48629308, pem=-0.20299491,
similarity=19.705585]
16: 25391 Russia
[pem_pos=1, final_scores=0.5703438,
final_local_scores=0.5838757, pem=-1.8802736,
similarity=31.738949]
17: 21148 Netherlands
[pem_pos=0, final_scores=0.5321602,
final_local_scores=0.52867883, pem=-0.24940795,
similarity=21.222864]
18: 17825936 Nathalie_Lancien
[pem_pos=0, final_scores=0.5090015,
final_local_scores=0.6460496, pem=0.0, similarity=23.369122]
19: 5843419 France
[pem_pos=0, final_scores=0.7729692,
```

```
final_local_scores=0.7569181, pem=-0.16240834,  
similarity=27.56086]
```

Sample: b'942_CYCLING'
CYCLING - WORLD TRACK CHAMPIONSHIP RESULTS . [1 1 MANCHESTER
1] , [2 2 England 2] 1996-08-30 Results at the world track
cycling championships on Friday : Women 's sprint semifinals
(best of three) : [3 3 Annett Neumann 3] ([4 4 Germany
4]) beat Magali Faure ([5 5 France 5]) 2-0 (12.341 and
12.348 seconds for the last 200 metres) [6 6 Félicia Ballanger
6] ([7 7 France 7]) beat [8 8 Tanya Dubnicoff 8] ([9 9 C
anada 9]) 2-0 , (12.130 / 12.124) Ride for third place :
Faure beat [10 10 Dubnicoff 10] 2-0 (12.112 / 12.246)
Final : [11 11 Ballanger 11] beat [12 12 Neumann 12] 2-0
(11.959 / 12.225) Women 's world points race championship
(24-km) : 1. Svetlana Samokhalova ([13 13 Russia 13]) 28
points (in 32 minutes 31.081 seconds) 2. Jane Quigley ([14
14 U.S. 14]) 18 points 3. Goulnara Fatkoullina ([15 15 Russ
ia 15]) 16 4. [16 16 Tatiana Stiakina 16] ([17 17 Ukraine 1
7]) 11 5. [18 18 Judith Arndt 18] ([19 19 Germany 19]) 11
6. [20 20 Tea Vikstedt-Nyman 20] ([21 21 Finland 21]) 5 One
lap behind : 7. Ingrid Haringa ([22 22 Netherlands 22]) 20
8. [23 23 Sally Boyden 23] ([24 24 Britain 24]) 9 9.
Agnieszka Godras ([25 25 Poland 25]) 8 10. [26 26 Nathalie L
ancien 26] ([27 27 France 27]) 8
1: 20206 Manchester
2: 9316 England
3: 6691583 Annett_Neumann
4: 11867 Germany
5: 5843419 France
6: 6519486 Félicia_Ballanger
7: 5843419 France
8: 11451462 Tanya_Dubnicoff
9: 5042916 Canada
10: 11451462 Tanya_Dubnicoff
11: 6519486 Félicia_Ballanger
12: 6691583 Annett_Neumann
13: 25391 Russia
14: 3434750 United_States
15: 25391 Russia
16: 18078979 Tatiana_Stiakina
17: 31750 Ukraine
18: 2196613 Judith_Arndt
19: 11867 Germany
20: 24910123 Tea_Vikstedt-Nyman
21: 10577 Finland
22: 21148 Netherlands
23: 14800477 Sally_Boyden_(cyclist)
24: 31717 United_Kingdom
25: 22936 Poland
26: 17825936 Nathalie_Lancien
27: 5843419 France

```
1: 20206 Manchester
    [pem_pos=0, final_scores=0.5552739,
final_local_scores=0.57738566, pem=-0.3209542,
similarity=23.065453]
2: 9316 England
    [pem_pos=0, final_scores=0.6658866,
final_local_scores=0.68757373, pem=-0.19053856,
similarity=25.644463]
3: 6691583 Annett_Neumann
    [pem_pos=0, final_scores=0.6134458,
final_local_scores=0.7543932, pem=0.0, similarity=26.602312]
4: 11867 Germany
    [pem_pos=0, final_scores=0.9261046,
final_local_scores=0.94292706, pem=-0.15529114,
similarity=33.073036]
5: 5843419 France
    [pem_pos=0, final_scores=0.79733974,
final_local_scores=0.8017749, pem=-0.16240834,
similarity=28.899477]
6: 6519486 Félicia_Ballanger
    [pem_pos=0, final_scores=0.85628253,
final_local_scores=1.0144694, pem=0.0, similarity=34.363506]
7: 5843419 France
    [pem_pos=0, final_scores=0.81448644,
final_local_scores=0.8205051, pem=-0.16240834,
similarity=29.458422]
8: 11451462 Tanya_Dubnicoff
    [pem_pos=0, final_scores=0.7741315,
final_local_scores=0.90371245, pem=0.0, similarity=31.058294]
9: 5042916 Canada
    [pem_pos=0, final_scores=0.84279543,
final_local_scores=0.8610295, pem=-0.11662793,
similarity=30.41879]
10: 11451462 Tanya_Dubnicoff
    [pem_pos=0, final_scores=0.46618932,
final_local_scores=0.5673324, pem=0.0, similarity=21.020042]
11: 6519486 Félicia_Ballanger
    [pem_pos=0, final_scores=0.6934719,
final_local_scores=0.83662355, pem=0.0, similarity=29.056229]
12: 6691583 Annett_Neumann
    [pem_pos=0, final_scores=0.54711753,
final_local_scores=0.6819397, pem=0.0, similarity=24.440155]
13: 25391 Russia
    [pem_pos=0, final_scores=0.7855231,
final_local_scores=0.8198, pem=-0.20299491,
similarity=29.658098]
14: 3434750 United_States
    [pem_pos=0, final_scores=0.50509983,
final_local_scores=0.5002255, pem=-0.14075686,
similarity=19.782898]
15: 25391 Russia
    [pem_pos=0, final_scores=0.76030475,
```

```
final_local_scores=0.7922528, pem=-0.20299491,
similarity=28.836033]
16: 18078979 Tatiana_Stajkina
    [pem_pos=0, final_scores=0.49918336,
final_local_scores=0.64684653, pem=0.0, similarity=23.392904]
17: 31750 Ukraine
    [pem_pos=0, final_scores=0.69431764,
final_local_scores=0.7062568, pem=-0.35407752,
similarity=27.091352]
18: 2196613 Judith_Arndt
    [pem_pos=0, final_scores=0.67091113,
final_local_scores=0.79179746, pem=0.0, similarity=27.718529]
19: 11867 Germany
    [pem_pos=0, final_scores=0.84741205,
final_local_scores=0.8569674, pem=-0.15529114,
similarity=30.507826]
20: 24910123 Tea_Vikstedt-Nyman
    [pem_pos=0, final_scores=0.574472,
final_local_scores=0.72885466, pem=0.0, similarity=25.840189]
21: 10577 Finland
    [pem_pos=0, final_scores=0.7809208,
final_local_scores=0.7941151, pem=-0.2445894,
similarity=29.117807]
22: 21148 Netherlands
    [pem_pos=0, final_scores=0.63243866,
final_local_scores=0.63556564, pem=-0.24940795,
similarity=24.412579]
23: 14800477 Sally_Boyden_(cyclist)
    [pem_pos=1, final_scores=0.5148099,
final_local_scores=0.65517086, pem=-0.14107856,
similarity=24.408525]
24: 31717 United_Kingdom
    [pem_pos=0, final_scores=0.5713002,
final_local_scores=0.5736386, pem=-0.61114657,
similarity=24.531744]
25: 22936 Poland
    [pem_pos=0, final_scores=0.553098,
final_local_scores=0.54276687, pem=-0.15333635,
similarity=21.120825]
26: 17825936 Nathalie_Lancien
    [pem_pos=0, final_scores=0.45867515,
final_local_scores=0.5778909, pem=0.0, similarity=21.335129]
27: 5843419 France
    [pem_pos=0, final_scores=0.44915843,
final_local_scores=0.42143977, pem=-0.16240834,
similarity=17.549517]
```

Sample: b'943_SOCCER'

SOCCER - [1 1 FRENCH 1] DEFENDER KOMBOUARE JOINS [2 2 ABERDEEN 2]. ABERDEEN , [3 3 Scotland 3] 1996-08-30 [4 4 French 4] central defender [5 5 Antoine Kombouare 5] has completed a 300,000 pounds sterling (\$ 467,000) move to [6 6 Aberdeen 6]

from [7 7 Swiss 7] club [8 8 Sion 8] , the [9 9 Scottish 9] premier division club said on Friday . Kombouare has signed a two-year contract and will make his debut against [10 10 Mort on 10] in the [11 11 Scottish League Cup 11] on Tuesday . But he will be ineligible for the rest of [12 12 Aberdeen 12] 's [13 13 UEFA Cup 13] campaign as he has already played for [14 1 4 Sion 14] in this season 's Cup Winners ' Cup . [15 15 Aberdeen 15] manager [16 16 Roy Aitken 16] said : " It 's unfortunate for us that Antoine cannot play in [17 17 Europe 1 7] but he will help us achieve things in domestic competition . " I have been watching him for several weeks now and have no doubts he brings real quality to the side . He has a great deal of experience and I 'm sure he will quickly establish himself in both the team and the affection of our fans . " The 32-year-old defender played seven seasons with [18 18 Nantes 18] and was with [19 19 Paris St Germain 19] for five seasons . He said former [20 20 PSG 20] team mate [21 21 David Ginola 21] , who now plays for [22 22 English 22] premier league [23 23 Newcastle 23] , was influential in his move to [24 24 Scotland 24] . " I 'm a very good friend of [25 25 David 25] and spoke to him recently about coming to [26 26 Aberdeen 26] and he was very positive about it , " Kombouare said . " He said I would really enjoy life there and that I would settle in in terms of football as well . That , and the fact he is only a few hours drive away , influenced my decision to come to [27 27 Aberdeen 27] . "

1: 250197 France_national_football_team
2: 237019 Aberdeen_F.C.
3: 26994 Scotland
4: 5843419 France
5: 8964356 Antoine_Kombouaré
6: 237019 Aberdeen_F.C.
7: 26748 Switzerland
8: 1061075 FC_Sion
9: 26994 Scotland
10: 367890 Greenock_Morton_F.C.
11: 1177084 Scottish_League_Cup
12: 237019 Aberdeen_F.C.
13: 232175 UEFA_Europa_League
14: 1061075 FC_Sion
15: 237019 Aberdeen_F.C.
16: 1317776 Roy_Aitken
17: 9239 Europe
18: 1135276 FC_Nantes
19: 357488 Paris_Saint-Germain_F.C.
20: 357488 Paris_Saint-Germain_F.C.
21: 1438258 David_Ginola
22: 9316 England
23: 57802 Newcastle_United_F.C.
24: 26994 Scotland
25: 1438258 David_Ginola
26: 237019 Aberdeen_F.C.

```
27: 237019 Aberdeen_F.C.
1: 250197 France_national_football_team
    [pem_pos=18, final_scores=0.48675638,
final_local_scores=0.46739966, pem=-5.6158586,
similarity=48.577763]
2: 237019 Aberdeen_F.C.
    [pem_pos=1, final_scores=0.65281874,
final_local_scores=0.61862063, pem=-1.4493513,
similarity=30.432384]
3: 26994 Scotland
    [pem_pos=1, final_scores=0.44611734,
final_local_scores=0.47777086, pem=-0.22905664,
similarity=19.592995]
4: 5843419 France
    [pem_pos=0, final_scores=0.6208133,
final_local_scores=0.6875699, pem=-0.80223274,
similarity=28.970837]
5: 8964356 Antoine_Kombouaré
    [pem_pos=0, final_scores=0.5696635,
final_local_scores=0.6332508, pem=0.0, similarity=22.987179]
6: 237019 Aberdeen_F.C.
    [pem_pos=1, final_scores=0.61908126,
final_local_scores=0.5817676, pem=-1.4493513,
similarity=29.332617]
7: 26748 Switzerland
    [pem_pos=0, final_scores=0.6451623,
final_local_scores=0.71656126, pem=-0.1392413,
similarity=26.230547]
8: 1061075 FC_Sion
    [pem_pos=1, final_scores=0.6241364,
final_local_scores=0.6338832, pem=-1.4381809,
similarity=30.827105]
9: 26994 Scotland
    [pem_pos=0, final_scores=0.49007672,
final_local_scores=0.52578986, pem=-0.47951394,
similarity=22.388002]
10: 367890 Greenock_Morton_F.C.
    [pem_pos=0, final_scores=0.6620564,
final_local_scores=0.65404505, pem=-1.8236359,
similarity=33.524937]
11: 1177084 Scottish_League_Cup
    [pem_pos=0, final_scores=0.35917205,
final_local_scores=0.34919265, pem=0.0, similarity=14.510318]
12: 237019 Aberdeen_F.C.
    [pem_pos=1, final_scores=0.80680984,
final_local_scores=0.7868326, pem=-1.4493513,
similarity=35.452164]
13: 232175 UEFA_Europa_League
    [pem_pos=0, final_scores=0.43042952,
final_local_scores=0.4170366, pem=-0.260892,
similarity=17.953688]
14: 1061075 FC_Sion
```

```
[pem_pos=1, final_scores=0.7837215,
final_local_scores=0.80820566, pem=-1.4381809,
similarity=36.029236]
15: 237019 Aberdeen_F.C.
    [pem_pos=1, final_scores=0.85555893,
final_local_scores=0.8400836, pem=-1.4493513,
similarity=37.041283]
16: 1317776 Roy_Aitken
    [pem_pos=0, final_scores=0.49981165,
final_local_scores=0.54859525, pem=0.0, similarity=20.46089]
17: 9239 Europe
    [pem_pos=0, final_scores=0.6446654,
final_local_scores=0.760073, pem=-0.086338006,
similarity=27.241327]
18: 1135276 FC_Nantes
    [pem_pos=1, final_scores=0.42272705,
final_local_scores=0.41404855, pem=-1.4553026,
similarity=24.35991]
19: 357488 Paris_Saint-Germain_F.C.
    [pem_pos=0, final_scores=0.4612831,
final_local_scores=0.44322616, pem=0.0, similarity=17.316465]
20: 357488 Paris_Saint-Germain_F.C.
    [pem_pos=0, final_scores=0.48986298,
final_local_scores=0.47444454, pem=-0.4638402,
similarity=20.770546]
21: 1438258 David_Ginola
    [pem_pos=0, final_scores=0.66453034,
final_local_scores=0.6935936, pem=0.0, similarity=24.78793]
22: 9316 England
    [pem_pos=1, final_scores=0.56522685,
final_local_scores=0.6241217, pem=-1.3527626,
similarity=30.071283]
23: 57802 Newcastle_United_F.C.
    [pem_pos=2, final_scores=0.7058756,
final_local_scores=0.6842928, pem=-2.4885974,
similarity=38.043755]
24: 26994 Scotland
    [pem_pos=0, final_scores=0.4464664,
final_local_scores=0.47815216, pem=-0.22905664,
similarity=19.604372]
25: 1438258 David_Ginola
    [pem_pos=0, final_scores=0.9058917,
final_local_scores=0.95724434, pem=0.0, similarity=32.655792]
26: 237019 Aberdeen_F.C.
    [pem_pos=1, final_scores=0.5277291,
final_local_scores=0.4819792, pem=-1.4493513,
similarity=26.35473]
27: 237019 Aberdeen_F.C.
    [pem_pos=1, final_scores=0.5495294,
final_local_scores=0.50579274, pem=-1.4493513,
similarity=27.065374]
```

Sample: b'944_MOTORCYCLING'

MOTORCYCLING - [1 1 SAN MARINO 1] GRAND PRIX PRACTICE
TIMES . [2 2 IMOLA 2] , [3 3 Italy 3] 1996-08-30 Practice
times set on Friday for Sunday 's [4 4 San Marino 4] 500cc
motorcycling Grand Prix : 1. [5 5 Michael Doohan 5] ([6 6 Au
stralia 6]) [7 7 Honda 7] one minute 50.250 2. [8 8 Jean-
Michel Bayle 8] ([9 9 France 9]) [10 10 Yamaha 10] 1:50.727
3. [11 11 Norifumi Abe 11] ([12 12 Japan 12]) [13 13 Yamaha
13] 1:50.858 4. [14 14 Luca Cadalora 14] ([15 15 Italy 15])
[16 16 Honda 16] 1:51.006 5. [17 17 Alex Criville 17] ([18 18
Spain 18]) [19 19 Honda 19] 1:51.075 6. [20 20 Scott Russell
20] ([21 21 United States 21]) [22 22 Suzuki 22] 1:51.287
7. [23 23 Tadayuki Okada 23] ([24 24 Japan 24]) [25 25 Honda
25] 1:51.528 8. [26 26 Carlos Checa 26] ([27 27 Spain 27])
[28 28 Honda 28] 1:51.588 9. Alexandre Barros ([29 29 Brazil 2
9]) [30 30 Honda 30] 1:51.784 10. [31 31 Shinichi Itoh 31] ([32 32 Japan 32]) [33 33 Honda 33] 1:51.857

1: 27248 San_Marino
2: 916593 Imola
3: 14532 Italy
4: 27248 San_Marino
5: 20108 Mick_Doohan
6: 4689264 Australia
7: 13729 Honda
8: 6042632 Jean-Michel_Bayle
9: 5843419 France
10: 332378 Yamaha_Motor_Company
11: 2857412 Norifumi_Abe
12: 15573 Japan
13: 332378 Yamaha_Motor_Company
14: 2035694 Luca_Cadalora
15: 14532 Italy
16: 13729 Honda
17: 6018692 Àlex_Crivillé
18: 26667 Spain
19: 13729 Honda
20: 8564874 Scott_Russell_(motorcyclist)
21: 3434750 United_States
22: 614614 Suzuki
23: 6018963 Tadayuki_Okada
24: 15573 Japan
25: 13729 Honda
26: 2993392 Carlos_Checa
27: 26667 Spain
28: 13729 Honda
29: 3383 Brazil
30: 13729 Honda
31: 12185077 Shinichi_Ito
32: 15573 Japan
33: 13729 Honda
1: 27248 San_Marino
[pem_pos=0, final_scores=0.41355342,

```
final_local_scores=0.4617597, pem=-0.78856933,
similarity=22.157906]
2: 916593 Imola
    [pem_pos=0, final_scores=0.39615804,
final_local_scores=0.458453, pem=-0.6584533,
similarity=21.351635]
3: 14532 Italy
    [pem_pos=0, final_scores=0.7158275,
final_local_scores=0.7592369, pem=-0.20612453,
similarity=27.867794]
4: 27248 San_Marino
    [pem_pos=0, final_scores=0.29903203,
final_local_scores=0.33666247, pem=-0.78856933,
similarity=18.424755]
5: 20108 Mick_Doohan
    [pem_pos=0, final_scores=0.5411502,
final_local_scores=0.58902967, pem=0.0, similarity=21.667534]
6: 4689264 Australia
    [pem_pos=0, final_scores=0.71014184,
final_local_scores=0.7777316, pem=-0.17168477,
similarity=28.232424]
7: 13729 Honda
    [pem_pos=0, final_scores=0.45533782,
final_local_scores=0.42434263, pem=-0.17204832,
similarity=17.688568]
8: 6042632 Jean-Michel_Bayle
    [pem_pos=0, final_scores=0.47111493,
final_local_scores=0.5737216, pem=0.0, similarity=21.210709]
9: 5843419 France
    [pem_pos=0, final_scores=0.72055644,
final_local_scores=0.7768375, pem=-0.16240834,
similarity=28.155294]
10: 332378 Yamaha_Motor_Company
    [pem_pos=1, final_scores=0.5460159,
final_local_scores=0.54962736, pem=-0.7066168,
similarity=24.334381]
11: 2857412 Norifumi_Abe
    [pem_pos=0, final_scores=0.46203166,
final_local_scores=0.5374223, pem=0.0, similarity=20.127466]
12: 15573 Japan
    [pem_pos=0, final_scores=0.73426795,
final_local_scores=0.78656137, pem=-0.13549899,
similarity=28.299137]
13: 332378 Yamaha_Motor_Company
    [pem_pos=1, final_scores=0.50710577,
final_local_scores=0.5071239, pem=-0.7066168,
similarity=23.065994]
14: 2035694 Luca_Cadalora
    [pem_pos=0, final_scores=0.74462664,
final_local_scores=0.8455315, pem=0.0, similarity=29.32206]
15: 14532 Italy
    [pem_pos=0, final_scores=0.7492356,
```

```
final_local_scores=0.79573, pem=-0.20612453,
similarity=28.956821]
16: 13729 Honda
    [pem_pos=0, final_scores=0.42841828,
final_local_scores=0.39493716, pem=-0.17204832,
similarity=16.81105]
17: 6018692 Àlex_Crivillé
    [pem_pos=0, final_scores=0.64523417,
final_local_scores=0.7408359, pem=0.0, similarity=26.197735]
18: 26667 Spain
    [pem_pos=0, final_scores=0.64253634,
final_local_scores=0.69117737, pem=-0.18857051,
similarity=25.7413]
19: 13729 Honda
    [pem_pos=0, final_scores=0.49242294,
final_local_scores=0.4648525, pem=-0.17204832,
similarity=18.897465]
20: 8564874 Scott_Russell_(motorcyclist)
    [pem_pos=1, final_scores=0.40235347,
final_local_scores=0.4704377, pem=-0.7074947,
similarity=21.975977]
21: 3434750 United_States
    [pem_pos=0, final_scores=0.5541406,
final_local_scores=0.59771293, pem=-0.073457584,
similarity=22.326134]
22: 614614 Suzuki
    [pem_pos=0, final_scores=0.4224642,
final_local_scores=0.4062879, pem=-0.1481934,
similarity=17.020054]
23: 6018963 Tadayuki_Okada
    [pem_pos=0, final_scores=0.42164582,
final_local_scores=0.52310437, pem=0.0, similarity=19.700191]
24: 15573 Japan
    [pem_pos=0, final_scores=0.76220244,
final_local_scores=0.81707555, pem=-0.13549899,
similarity=29.209742]
25: 13729 Honda
    [pem_pos=0, final_scores=0.51913625,
final_local_scores=0.4940328, pem=-0.17204832,
similarity=19.768263]
26: 2993392 Carlos_Checa
    [pem_pos=0, final_scores=0.45739204,
final_local_scores=0.515379, pem=0.0, similarity=19.46965]
27: 26667 Spain
    [pem_pos=0, final_scores=0.62152535,
final_local_scores=0.66822606, pem=-0.18857051,
similarity=25.056387]
28: 13729 Honda
    [pem_pos=0, final_scores=0.51732844,
final_local_scores=0.49205798, pem=-0.17204832,
similarity=19.70933]
29: 3383 Brazil
```

```

    [pem_pos=0, final_scores=0.75347334,
final_local_scores=0.8342078, pem=-0.21313342,
similarity=30.14319]
30: 13729 Honda
    [pem_pos=0, final_scores=0.50342375,
final_local_scores=0.4768693, pem=-0.17204832,
similarity=19.25607]
31: 12185077 Shinichi_Ito
    [pem_pos=0, final_scores=0.58412653,
final_local_scores=0.66325563, pem=0.0, similarity=23.882584]
32: 15573 Japan
    [pem_pos=0, final_scores=0.740331,
final_local_scores=0.79318434, pem=-0.13549899,
similarity=28.49678]
33: 13729 Honda
    [pem_pos=0, final_scores=0.40976304,
final_local_scores=0.3745591, pem=-0.17204832,
similarity=16.202929]

```

Sample: b'945_GOLF'

GOLF - [1 1 BRITISH MASTERS 1] THIRD ROUND SCORES . [2 2 NOR THAMPTON 2] , [3 3 England 3] 1996-08-30 Leading scores after the third round of the [4 4 British Masters 4] on Friday :
211 [5 5 Robert Allenby 5] ([6 6 Australia 6]) 69 71 71 212
[7 7 Pedro Linhart 7] ([8 8 Spain 8]) 72 73 67 216 [9 9 Miguel Angel Martin 9] ([10 10 Spain 10]) 75 70 71 , [11 11 Costantino Rocca 11] ([12 12 Italy 12]) 71 73 72 217 Antoine Lebouc ([13 13 France 13]) 74 73 70 , [14 14 Ian Woosnam 14] 70 76 71 , Francisco Cea ([15 15 Spain 15]) 70 71 76 , Gavin Levenson ([16 16 South Africa 16]) 66 75 76 218 [17 17 Stephen McAllister 17] 73 76 69 , [18 18 Joakim Haeggman 18] ([19 19 Swe 19]) 71 77 70 , [20 20 Jose Cceres 20] ([21 21 Argentina 21]) 69 78 71 , [22 22 Paul Eales 22] 75 71 72 , [23 23 Klas Eriksson 23] ([24 24 Sweden 24]) 71 75 72 , Mike Clayton ([25 25 Australia 25]) 69 76 73 , [26 26 Mark Roe 26] 69 71 78 219 [27 27 Eamonn Darcy 27] ([28 28 Ireland 28]) 74 76 69 , [29 29 Bob May 29] ([30 30 U.S. 30]) 74 75 70 , [31 31 Paul Lawrie 31] 72 75 72 , [32 32 Miguel Angel Jimenez 32] ([33 33 Spain 33]) 74 72 73 , [34 34 Peter Mitchell 34] 74 71 75 , [35 35 Philip Walton 35] ([36 36 Ireland 36]) 71 74 74 , Peter O'Malley ([37 37 Australia 37]) 71 73 75 220 [38 38 Barry Lane 38] 73 77 70 , [39 39 Wayne Riley 39] ([40 40 Australia 40]) 71 78 71 , Martin Gates 71 77 72 , [41 41 Bradley Hughes 41] ([42 42 Australia 42]) 73 75 72 , [43 43 Peter Hedblom 43] ([44 44 Sweden 44]) 70 75 75 , [45 45 Retief Goosen 45] ([46 46 South Africa 46]) 71 74 75 , [47 47 David Gilford 47] 69 74 77 .

1: 1929870 British_Masters
2: 63734 Northampton
3: 9316 England
4: 1929870 British_Masters
5: 1983254 Robert_Allenby

```
6: 4689264 Australia
7: 5285076 Pedro_Linhart
8: 26667 Spain
9: 5022198 Miguel_Ángel_Martín_(golfer)
10: 26667 Spain
11: 2157151 Costantino_Rocca
12: 14532 Italy
13: 5843419 France
14: 1034991 Ian_Woosnam
15: 26667 Spain
16: 17416221 South_Africa
17: 5266169 Stephen_McAllister
18: 2032221 Joakim_Haeggman
19: 5058739 Sweden
20: 2156041 José_Cóceres
21: 18951905 Argentina
22: 4435993 Paul_Eales
23: 28186295 Klas_Eriksson
24: 5058739 Sweden
25: 4689264 Australia
26: 4353394 Mark_Roe
27: 2770292 Eamonn_Darcy
28: 14560 Republic_of_Ireland
29: 3528782 Bob_May_(golfer)
30: 3434750 United_States
31: 1564436 Paul_Lawrie
32: 1569625 Miguel_Ángel_Jiménez
33: 26667 Spain
34: 5000237 Peter_Mitchell_(golfer)
35: 5179238 Philip_Walton
36: 14560 Republic_of_Ireland
37: 4689264 Australia
38: 2167839 Barry_Lane
39: 5243170 Wayne_Riley
40: 4689264 Australia
41: 9923018 Bradley_Hughes_(golfer)
42: 4689264 Australia
43: 4377532 Peter_Hedblom
44: 5058739 Sweden
45: 630525 Retief_Goosen
46: 17416221 South_Africa
47: 2028738 David_Gilford
1: 1929870 British_Masters
    [pem_pos=0, final_scores=0.4001149,
final_local_scores=0.47164392, pem=0.0, similarity=18.164509]
2: 63734 Northampton
    [pem_pos=0, final_scores=0.4250688,
final_local_scores=0.47950572, pem=-0.5110059,
similarity=21.178051]
3: 9316 England
    [pem_pos=0, final_scores=0.5712015,
final_local_scores=0.59702086, pem=-0.19053856,
```

```
similarity=22.942184]
4: 1929870 British_Masters
    [pem_pos=0, final_scores=0.3987772,
final_local_scores=0.47018266, pem=0.0, similarity=18.120901]
5: 1983254 Robert_Allenby
    [pem_pos=0, final_scores=0.72048414,
final_local_scores=0.8021575, pem=0.0, similarity=28.027693]
6: 4689264 Australia
    [pem_pos=0, final_scores=0.6094375,
final_local_scores=0.63740975, pem=-0.17168477,
similarity=24.04494]
7: 5285076 Pedro_Linhart
    [pem_pos=0, final_scores=0.6908662,
final_local_scores=0.7853167, pem=0.0, similarity=27.52513]
8: 26667 Spain
    [pem_pos=0, final_scores=0.5994812,
final_local_scores=0.6180046, pem=-0.18857051,
similarity=23.557678]
9: 5022198 Miguel_Ángel_Martín_(golfer)
    [pem_pos=0, final_scores=0.7021223,
final_local_scores=0.78606194, pem=0.0, similarity=27.547369]
10: 26667 Spain
    [pem_pos=0, final_scores=0.6263861,
final_local_scores=0.6473941, pem=-0.18857051,
similarity=24.434721]
11: 2157151 Costantino_Rocca
    [pem_pos=0, final_scores=0.6341695,
final_local_scores=0.7005214, pem=0.0, similarity=24.994669]
12: 14532 Italy
    [pem_pos=0, final_scores=0.64840233,
final_local_scores=0.69972813, pem=-0.20612453,
similarity=26.091934]
13: 5843419 France
    [pem_pos=0, final_scores=0.66842407,
final_local_scores=0.7094125, pem=-0.16240834,
similarity=26.143198]
14: 1034991 Ian_Woosnam
    [pem_pos=0, final_scores=0.6549379,
final_local_scores=0.7005293, pem=0.0, similarity=24.994904]
15: 26667 Spain
    [pem_pos=0, final_scores=0.6220155,
final_local_scores=0.64261985, pem=-0.18857051,
similarity=24.292246]
16: 17416221 South_Africa
    [pem_pos=0, final_scores=0.5421267,
final_local_scores=0.5841313, pem=-0.26251382,
similarity=22.948948]
17: 5266169 Stephen_McAllister
    [pem_pos=0, final_scores=0.7034499,
final_local_scores=0.7927932, pem=0.0, similarity=27.748243]
18: 2032221 Joakim_Haeggman
    [pem_pos=0, final_scores=0.5993992,
```

```
final_local_scores=0.6740605, pem=0.0, similarity=24.205023]
19: 5058739 Sweden
    [pem_pos=0, final_scores=0.59699494,
final_local_scores=0.61977345, pem=-0.0023738882,
similarity=22.597898]
20: 2156041 José_Cóceres
    [pem_pos=0, final_scores=0.7488958,
final_local_scores=0.8190679, pem=0.0, similarity=28.532333]
21: 18951905 Argentina
    [pem_pos=0, final_scores=0.6677646,
final_local_scores=0.7028326, pem=-0.30003807,
similarity=26.695292]
22: 4435993 Paul_Eales
    [pem_pos=0, final_scores=0.62822723,
final_local_scores=0.7064993, pem=0.0, similarity=25.173061]
23: 28186295 Klas_Eriksson
    [pem_pos=0, final_scores=0.6222521,
final_local_scores=0.7353338, pem=0.0, similarity=26.03354]
24: 5058739 Sweden
    [pem_pos=0, final_scores=0.75930625,
final_local_scores=0.79707384, pem=-0.2135508,
similarity=29.037308]
25: 4689264 Australia
    [pem_pos=0, final_scores=0.753135,
final_local_scores=0.7943775, pem=-0.17168477,
similarity=28.729172]
26: 4353394 Mark_Roe
    [pem_pos=0, final_scores=0.6841171,
final_local_scores=0.7737244, pem=0.0, similarity=27.179192]
27: 2770292 Eamonn_Darcy
    [pem_pos=0, final_scores=0.5345819,
final_local_scores=0.60835904, pem=0.0, similarity=22.24436]
28: 14560 Republic_of_Ireland
    [pem_pos=1, final_scores=0.5792101,
final_local_scores=0.6163842, pem=-1.5040284,
similarity=30.662987]
29: 3528782 Bob_May_(golfer)
    [pem_pos=1, final_scores=0.5415143,
final_local_scores=0.5970135, pem=-0.48494384,
similarity=24.542986]
30: 3434750 United_States
    [pem_pos=0, final_scores=0.4668398,
final_local_scores=0.49163932, pem=-0.14075686,
similarity=19.526669]
31: 1564436 Paul_Lawrie
    [pem_pos=0, final_scores=0.72581416,
final_local_scores=0.8031104, pem=0.0, similarity=28.05613]
32: 1569625 Miguel_Ángel_Jiménez
    [pem_pos=0, final_scores=0.4925481,
final_local_scores=0.5515036, pem=0.0, similarity=20.54768]
33: 26667 Spain
    [pem_pos=0, final_scores=0.5929177,
```

```
final_local_scores=0.6108349, pem=-0.18857051,  
similarity=23.34372]  
34: 5000237 Peter_Mitchell_(golfer)  
    [pem_pos=3, final_scores=0.5537769,  
final_local_scores=0.62351894, pem=-1.2033675,  
similarity=29.240862]  
35: 5179238 Philip_Walton  
    [pem_pos=0, final_scores=0.815497,  
final_local_scores=0.8868707, pem=0.0, similarity=30.555704]  
36: 14560 Republic_of_Ireland  
    [pem_pos=1, final_scores=0.5932881,  
final_local_scores=0.6317623, pem=-1.5040284,  
similarity=31.1219]  
37: 4689264 Australia  
    [pem_pos=0, final_scores=0.6760189,  
final_local_scores=0.7101398, pem=-0.17168477,  
similarity=26.215351]  
38: 2167839 Barry_Lane  
    [pem_pos=0, final_scores=0.70798016,  
final_local_scores=0.79184926, pem=0.0, similarity=27.720074]  
39: 5243170 Wayne_Riley  
    [pem_pos=0, final_scores=0.6417139,  
final_local_scores=0.7049529, pem=0.0, similarity=25.126913]  
40: 4689264 Australia  
    [pem_pos=0, final_scores=0.7040455,  
final_local_scores=0.74075454, pem=-0.17168477,  
similarity=27.128956]  
41: 9923018 Bradley_Hughes_(golfer)  
    [pem_pos=0, final_scores=0.61688393,  
final_local_scores=0.7028628, pem=0.0, similarity=25.06454]  
42: 4689264 Australia  
    [pem_pos=0, final_scores=0.720221,  
final_local_scores=0.75842386, pem=-0.17168477,  
similarity=27.656242]  
43: 4377532 Peter_Hedblom  
    [pem_pos=0, final_scores=0.5982436,  
final_local_scores=0.6711582, pem=0.0, similarity=24.118412]  
44: 5058739 Sweden  
    [pem_pos=0, final_scores=0.76459175,  
final_local_scores=0.80284756, pem=-0.2135508,  
similarity=29.209608]  
45: 630525 Retief_Goosen  
    [pem_pos=0, final_scores=0.69144034,  
final_local_scores=0.77756596, pem=0.0, similarity=27.293833]  
46: 17416221 South_Africa  
    [pem_pos=0, final_scores=0.5796613,  
final_local_scores=0.6251322, pem=-0.26251382,  
similarity=24.172497]  
47: 2028738 David_Gilford  
    [pem_pos=0, final_scores=0.7679083,  
final_local_scores=0.8749004, pem=0.0, similarity=30.198484]
```

Sample: b'946_SOCCER'
SOCCER - [1 1 ENGLISH 1] SOCCER RESULTS . [2 2 LONDON 2]
1996-08-30 Results of [3 3 English 3] league matches on
Friday : Division two 1 Plymouth 2 [4 4 Preston 4] 1 Division
three [5 5 Swansea 5] 1 2 Lincoln 2
1: 9316 England
2: 17867 London
3: 9316 England
4: 429092 Preston_North_End_F.C.
5: 451169 Swansea_City_A.F.C.
1: 23857 Plymouth_Argyle_F.C.
[winner: 23508 Plymouth, pem_pos=0,
final_scores=-0.5587237, final_local_scores=-0.6596861,
pem=-0.45987546, similarity=-13.095761]
RECALL MISS: 23857 Plymouth_Argyle_F.C.]
2: 451163 Lincoln_City_F.C.
[winner: 62705 Lincoln_Motor_Company, pem_pos=2,
final_scores=-0.57893944, final_local_scores=-0.5378492,
pem=-2.4392905, similarity=1.3044612]
RECALL MISS: 451163 Lincoln_City_F.C.]
1: 9316 England
[pem_pos=1, final_scores=0.5085725,
final_local_scores=0.48118144, pem=-1.3527626,
similarity=25.80566]
2: 17867 London
[pem_pos=0, final_scores=0.67164874,
final_local_scores=0.6878469, pem=-0.09086737,
similarity=25.110586]
3: 9316 England
[pem_pos=1, final_scores=0.7226973,
final_local_scores=0.7150803, pem=-1.3527626,
similarity=32.785667]
4: 429092 Preston_North_End_F.C.
[pem_pos=4, final_scores=0.8023252,
final_local_scores=0.8509508, pem=-3.194944,
similarity=46.858383]
5: 451169 Swansea_City_A.F.C.
[pem_pos=3, final_scores=0.6749826,
final_local_scores=0.71481663, pem=-2.5988693,
similarity=39.55432]